

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/18, C07K 14/505, 14/52, A61K 38/18, C12N 5/10, 5/08	A1	(11) International Publication Number: WO 98/18926 (43) International Publication Date: 7 May 1998 (07.05.98)
(21) International Application Number: PCT/US97/18703 (22) International Filing Date: 23 October 1997 (23.10.97) (30) Priority Data: 60/034,044 25 October 1996 (25.10.96) US (71) Applicant (for all designated States except US): G.D. SEARLE & CO. [US/US]; Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): MCWHERTER, Charles, A. [US/US]; 16564 Thunderhead Canyon Court, Ellisville, MO 63011 (US). FENG, Yiqing [US/US]; 423 Mission Court, St. Louis, MO 63130 (US). SUMMERS, Nccna [US/US]; 1203 Saddlemaker, St. Charles, MO 63304 (US). (74) Agents: BENNETT, Dennis, A. et al.; G.D. Searle & Co., Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US).	(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: CIRCULARLY PERMUTED ERYTHROPOIETIN RECEPTOR AGONISTS (57) Abstract Disclosed are novel Erythropoietin receptor agonist proteins, DNAs which encode the Erythropoietin receptor agonist proteins, methods of making the Erythropoietin receptor agonist proteins and methods of using the Erythropoietin receptor agonist proteins.		

ATTORNEY DOCKET NUMBER: 10165-042-999
SERIAL NUMBER: 10/573,905
REFERENCE: **B45**

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

CIRCULARLY PERMUTED ERYTHROPOIETIN RECEPTOR AGONISTS

The present application claims priority under Title 35,
United States Code, §119 of United States Provisional
5 application Serial No. 60/034,044, filed October 25,
1996.

FIELD OF THE INVENTION

The present invention relates to human
10 Erythropoietin (EPO) receptor agonists. These EPO
receptor agonists retain one or more activities of
native EPO and may also show improved hematopoietic
cell-stimulating activity and/or an improved activity
profile which may include reduction of undesirable
15 biological activities associated with native EPO and/or
have improved physical properties which may include
increased solubility, stability and refold efficiency.

BACKGROUND OF THE INVENTION

20 Colony stimulating factors which stimulate the
differentiation and/or proliferation of bone marrow
cells have generated much interest because of their
therapeutic potential for restoring depressed levels of
hematopoietic stem cell-derived cells.

25 Erythropoietin is a naturally-occurring
glycoprotein hormone with a molecular weight that was
first reported to be approximately 39,000 daltons (T.
Miyaki et al., *J. Biol. Chem.* **252**:5558-5564 (1977)).
30 The mature hormone is 166 amino acids long and the
"prepro" form of the hormone, with its leader peptide,
is 193 amino acids long (F. Lin, U.S. Patent No.
4,703,008). The mature hormone has a molecular weight,
calculated from its amino acid sequence, of 18,399
35 daltons (K. Jacobs et al., *Nature* **313**:806-810 (1985);
J. K. Browne et al., *Cold Spring Harbor Symp. Quant.
Biol.* **5**:1693-702 (1986)).

2

The first mutant erythropoietins (i.e., erythropoietin analogs), prepared by making amino acid substitutions and deletions, have demonstrated reduced or unimproved activity. As described in U.S. Patent NO. 4,703,008, replacement of the tyrosine residues at positions 15, 40 and 145 with phenylalanine residues, replacement of the cysteine residue at position 7 with an histidine, substitution of the proline at position 2 with an asparagine, deletion of residues 2-6, deletion of residues 163-166, and deletion of residues 27-55 does not result in an apparent increase in biological activity. The Cys'-to-His' mutation eliminates biological activity. A series of mutant erythropoietins with a single amino acid substitution at asparagine residues 24, 38 or 83 show severely reduced activity (substitution at position 24) or exhibit rapid intracellular degradation and apparent lack of secretion (substitution at residue 38 or 183). Elimination of the O-linked glycosylation site at serine126 results in rapid degradation or lack of secretion of the erythropoietin analog (S. Dube et al., *J. Biol. Chem.* **33**:17516-17521 (1988)). These authors conclude that glycosylation sites at residues 38, 83 and 126 are required for proper secretion and that glycosylation sites located at residues 24 and 38 may be involved in the biological activity of mature erythropoietin.

Deglycosylated erythropoietin is fully active in *in vitro* bioassays (M. S. Dorsdal et al., *Endocrinology* **116**:2293-2299 (1985); U.S. Patent No. 4,703,008; E. Tsuda et al., *Eur J. Biochem.* **266**:20434-20439 (1991)). However, glycosylation of erythropoietin is widely accepted to play a critical role in the *in vivo* activity of the hormone (P. H. Lowy et al., *Nature* **185**:102-105 (1960); E. Goldwasser and C. K. H. Kung, *Ann. N.Y. Acad. Science* **149**:49-53 (1968); W. A. Lukowsky and R.

- H.. Painter, *Can. J. Biochem.* :909-917 (1972); D.W. Briggs et al., *Amer. J. Phys.* **201**:1385-1388 (1974); J.C. Schooley, *Exp. Hematol.* **13**:994-998; N. Imai et al., *Eur. J. Biochem.* **194**:457-462 (1990); M.S. Dordal et al.,
5 *Endocrinology* **116**:2293-2299 (1985); E. Tsuda et al.,
Eur. J. Biochem. **188**:405-411 (1990); U.S. Patent No.
4,703,008; J.K. Brown et al., *Cold Spring Harbor*
Symposia on Quant. Biol. **51**:693-702 (1986); and K.
Yamaguchi et al., *J. Biol. Chem.* **266**:20434-20439 (1991).
10 The lack if *in vivo* biological activity of
deglycosylated analogs of erythropoietin is attributed
to a rapid clearance of the deglycosylated hormone from
the circulation of treated animals. This view is
supported by direct comparison of the plasma half-life
15 of glycosylated and deglycosylated erythropoietin (J.C.
Spivak and B.B. Hoyans, *Blood* **73**:90-99 (1989), and M.N.
Fukuda, et al., *Blood* **73**:84-89 (1989).

20 Oligonucleotide-directed mutagenesis of
erythropoietin glycosylation sites has effectively
probed the function of glycosylation but has failed, as
yet, to provide insight into an effective strategy for
significantly improving the characteristics of the
hormone for therapeutic applications.

25

A series of single amino acid substitution or
deletion mutants have been constructed, involving amino
acid residues 15, 24, 49, 76, 78, 83, 143, 145, 160,
162, 163, 164, 165 and 166. In these mutants are altered
30 the carboxy terminus, the glycosylation sites, and the
tyrosine residues of erythropoietin. The mutants have
been administered to animals while monitoring
hemoglobin, hematocrit and reticulocyte levels (EP No. 0
409 113). While many of these mutants retain *in vivo*
35 biological activity, none show a significant increase in
their ability to raise hemoglobin, hematocrit or

4

reticulocyte (the immediate precursor of an erythrocyte) levels when compared to native erythropoietin.

Another set of mutants has been constructed to probe the function of residues 99-119 (domain 1) and residues 111-129 (domain 2) (Y. Chern et al., *Eur. J. Biochem.* **202**:225-230 (1991)). The domain 1 mutants are rapidly degraded and inactive in an *in vitro* bioassay while the domain 2 mutants, at best, retain *in vitro* activity. These mutants also show no enhanced *in vivo* biological activity as compared to wild-type, human erythropoietin. These authors conclude that residues 99-119 play a critical role in the structure of erythropoietin.

The human erythropoietin molecule contains two disulfide bridges, one linking the cysteine residues at positions 7 and 161, and a second connecting cysteines at positions 29 and 33 (P.H. Lai et al., *J. Biol. Chem.* **261**:3116-3121 (1986)). Oligonucleotide-directed mutagenesis has been used to probe the function of the disulfide bridge linking cysteines 29 and 33 in human erythropoietin. The cysteine at position 33 has been converted to a proline residue, which, mimics the structure of murine erythropoietin at this residue. The resulting mutant has greatly reduced *in vitro* activity. The loss of activity is so severe that the authors conclude that the disulfide bridge between residues 29 and 33 is essential for erythropoietin function (F.K. Lin, *Molecular and Cellular Aspects of Erythropoietin and Erythropoiesis*, pp. 23-36, ed. I.N. Rich, Springer-Verlag, Berlin (1987)).

U.S. Patent No. 4,703,008 by Lin, F-K. (hereinafter referred to as "the '008 patent") speculates about a wide variety of modifications of EPO, including addition, deletion, and substitution analogs of EPO.

5

The '008 patent does not indicate that any of the suggested modifications would increase biological activity *per se*, although it is stated that deletion of glycosylation sites might increase the activity of EPO produced in yeast (See the '008 patent at column 37, lines 25-28). Also, the '008 patent speculates that EPO analogs which have one or more tyrosine residues replaced with phenylalanine may exhibit an increased or decreased receptor binding affinity.

10

Australian Patent Application No. AU-A-59145/90 by Fibi, M et al. also discusses a number of modified EPO proteins (EPO muteins). It is generally speculated that the alteration of amino acids 10-55, 70-85, and 130-166 of EPO. In particular, additions of positively charged basic amino acids in the carboxyl terminal region are purported to increase the biological activity of EPO.

U.S. Patent No. 4,835,260 by Shoemaker, C.B. discusses modified EPO proteins with amino acid substitutions of the methionine at position 54 and asparagine at position 38. Such EPO muteins are thought to have improved stability but are not proposed to exhibit any increase in biological activity relative to wild type EPO.

WO 91/05867 discloses analogs of human erythropoietin having a greater number of sites for carbohydrate attachment than human erythropoietin, such as EPO (Asn⁶⁹), EPO (Asn¹²⁵, Ser¹²⁷), EPO (Thr¹²⁵), and EPO (Pro¹²⁴, Thr¹²⁵).

WO 94 /24160 discloses erythropoietin muteins which have enhanced activity, specifically amino acid substitutions at positions 20, 49, 73, 140, 143, 146, 147 and 154.

WO 94/25055 discloses erythropoietin analogs, including EPO (X³³, Cys¹³⁹, des-Arg¹⁶⁶) and EPO (Cys¹³⁹, des-Arg¹⁶⁶).

5

Rearrangement of Protein Sequences

In evolution, rearrangements of DNA sequences serve an important role in generating a diversity of protein structure and function. Gene duplication and exon shuffling provide an important mechanism to rapidly generate diversity and thereby provide organisms with a competitive advantage, especially since the basal mutation rate is low (Doolittle, *Protein Science* 1:191-200, 1992).

The development of recombinant DNA methods has made it possible to study the effects of sequence transposition on protein folding, structure and function. The approach used in creating new sequences resembles that of naturally occurring pairs of proteins that are related by linear reorganization of their amino acid sequences (Cunningham, et al., *Proc. Natl. Acad. Sci. U.S.A.* **76**:3218-3222, 1979; Teather & Erfle, *J. Bacteriol.* **172**: 3837-3841, 1990; Schimming et al., *Eur. J. Biochem.* **204**: 13-19, 1992; Yamiuchi and Minamikawa, *FEBS Lett.* **260**:127-130, 1991; MacGregor et al., *FEBS Lett.* **378**:263-266, 1996). The first in vitro application of this type of rearrangement to proteins was described by Goldenberg and Creighton (*J. Mol. Biol.* **165**:407-413, 1983). A new N-terminus is selected at an internal site (breakpoint) of the original sequence, the new sequence having the same order of amino acids as the original from the breakpoint until it reaches an amino acid that is at or near the original C-terminus. At this point the new sequence is joined, either directly or through an additional portion of sequence (linker), to an amino acid that is at or near the original N-

7

terminus, and the new sequence continues with the same sequence as the original until it reaches a point that is at or near the amino acid that was N-terminal to the breakpoint site of the original sequence, this residue
 5 forming the new C-terminus of the chain.

This approach has been applied to proteins which range in size from 58 to 462 amino acids (Goldenberg & Creighton, *J. Mol. Biol.* **165**:407-413, 1983; Li & Coffino, *Mol. Cell. Biol.* **13**:2377-2383, 1993). The
 10 proteins examined have represented a broad range of structural classes, including proteins that contain predominantly α -helix (interleukin-4; Kreitman et al., *Cytokine* **7**:311-318, 1995), β -sheet (interleukin-1; Horlick et al., *Protein Eng.* **5**:427-431, 1992), or
 15 mixtures of the two (yeast phosphoribosyl anthranilate isomerase; Luger et al., *Science* **243**:206-210, 1989). Broad categories of protein function are represented in these sequence reorganization studies:

20 **Enzymes**

- | | | |
|----|------------------------------------|--|
| 25 | T4 lysozyme | Zhang et al., <i>Biochemistry</i> 32 :12311-12318 (1993); Zhang et al., <i>Nature Struct. Biol.</i> 1 :434-438 (1995) |
| 30 | dihydrofolate reductase | Buchwalder et al., <i>Biochemistry</i> 31 :1621-1630 (1994); Protasova et al., <i>Prot. Eng.</i> 7 :1373-1377 (1995) |
| 35 | ribonuclease T1 | Mullins et al., <i>J. Am. Chem. Soc.</i> 116 :5529-5533 (1994); Garrett et al., <i>Protein Science</i> 5 :204-211 (1996) |
| | <i>Bacillus</i> β -glucanase | Hahn et al., <i>Proc. Natl. Acad. Sci. U.S.A.</i> 91 :10417-10421 (1994) |

8

- aspartate
transcarbamoylase Yang & Schachman, *Proc. Natl. Acad. Sci. U.S.A.* **90**:11980-11984 (1993)
- phosphoribosyl
5 anthranilate Luger et al., *Science* **243**:206-210
isomerase (1989); Luger et al., *Prot. Eng.*
3:249-258 (1990)
- pepsin/pepsinogen Lin et al., *Protein Science* **4**:159-
166 (1995)
- 10 glyceraldehyde-3- Vignais et al., *Protein Science*
phosphate dehydro- **4**:994-1000 (1995)
genase
- 15 ornithine Li & Coffino, *Mol. Cell. Biol.*
decarboxylase **13**:2377-2383 (1993)
- yeast
phosphoglycerate Ritco-Vonsovici et al., *Biochemistry*
20 dehydrogenase **34**:16543-16551 (1995)

Enzyme Inhibitor

- basic pancreatic Goldenberg & Creighton, *J. Mol.*
25 trypsin inhibitor *Biol.* **165**:407-413 (1983)

Cytokines

- interleukin-1 β Horlick et al., *Protein Eng.* **5**:427-
30 431 (1992)
- interleukin-4 Kreitman et al., *Cytokine* **7**:311-
318 (1995)

- 35 **Tyrosine Kinase**
Recognition Domain

9

α -spectrin SH3 domain Viguera, et al., *J. Mol. Biol.* **247**:670-681 (1995)

Transmembrane5 **Protein**

omp A Koebnik & Krämer, *J. Mol. Biol.* **250**:617-626 (1995)

10 **Chimeric Protein**

interleukin-4-*Pseudomonas* exotoxin fusion molecule Kreitman et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**:6889-6893 (1994).

The results of these studies have been highly variable. In many cases substantially lower activity, solubility or thermodynamic stability were observed (*E. coli* dihydrofolate reductase, aspartate transcarbamoylase, phosphoribosyl anthranilate isomerase, glyceraldehyde-3-phosphate dehydrogenase, ornithine decarboxylase, omp A, yeast phosphoglycerate dehydrogenase). In other cases, the sequence rearranged protein appeared to have many nearly identical properties as its natural counterpart (basic pancreatic trypsin inhibitor, T4 lysozyme, ribonuclease T1, *Bacillus* β -glucanase, interleukin-1 β , α -spectrin SH3 domain, pepsinogen, interleukin-4). In exceptional cases, an unexpected improvement over some properties of the natural sequence was observed, e.g., the solubility and refolding rate for rearranged α -spectrin SH3 domain sequences, and the receptor affinity and anti-tumor activity of transposed interleukin-4-*Pseudomonas* exotoxin fusion molecule (Kreitman et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**:6889-6893, 1994; Kreitman et al., *Cancer Res.* **55**:3357-3363, 1995).

The primary motivation for these types of studies has been to study the role of short-range and long-range

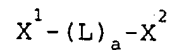
10

interactions in protein folding and stability. Sequence rearrangements of this type convert a subset of interactions that are long-range in the original sequence into short-range interactions in the new sequence, and vice versa. The fact that many of these sequence rearrangements are able to attain a conformation with at least some activity is persuasive evidence that protein folding occurs by multiple folding pathways (Viguera, et al., *J. Mol. Biol.* **247**:670-681, 1995). In the case of the SH3 domain of α -spectrin, choosing new termini at locations that corresponded to β -hairpin turns resulted in proteins with slightly less stability, but which were nevertheless able to fold.

The positions of the internal breakpoints used in the studies cited here are found exclusively on the surface of proteins, and are distributed throughout the linear sequence without any obvious bias towards the ends or the middle (the variation in the relative distance from the original N-terminus to the breakpoint is ca. 10 to 80% of the total sequence length). The linkers connecting the original N- and C-termini in these studies have ranged from 0 to 9 residues. In one case (Yang & Schachman, *Proc. Natl. Acad. Sci. U.S.A.* **90**:11980-11984, 1993), a portion of sequence has been deleted from the original C-terminal segment, and the connection made from the truncated C-terminus to the original N-terminus. Flexible hydrophilic residues such as Gly and Ser are frequently used in the linkers. Viguera, et al. (*J. Mol. Biol.* **247**:670-681, 1995) compared joining the original N- and C-termini with 3- or 4-residue linkers; the 3-residue linker was less thermodynamically stable. Protasova et al. (*Protein Eng.* **7**:1373-1377, 1994) used 3- or 5-residue linkers in connecting the original N-termini of *E. coli* dihydrofolate reductase; only the 3-residue linker produced protein in good yield.

//
Summary of the Invention

The modified human EPO receptor agonists of the
 5 present invention can be represented by the Formula:



wherein;

- 10 a is 0 or 1;
 X^1 is a peptide comprising an amino acid
 sequence corresponding to the sequence of residues n+1
 through J;
 X^2 is a peptide comprising an amino acid
 15 sequence corresponding to the sequence of residues 1
 through n;
 n is an integer ranging from 1 to J-1; and
 L is a linker.

- 20 In the formula above the constituent amino acids
 residues of human EPO are numbered sequentially 1
 through J from the amino to the carboxyl terminus. A
 pair of adjacent amino acids within this protein may be
 numbered n and n+1 respectively where n is an integer
 25 ranging from 1 to J-1. The residue n+1 becomes the new
 N-terminus of the new EPO receptor agonist and the
 residue n becomes the new C-terminus of the new EPO
 receptor agonist.

- 30 The present invention relates to novel EPO receptor
 agonists polypeptides comprising a modified EPO amino
 acid sequence of the following formula:

- 35 AlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys
 10 20
 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr
 30 40
 40 ValProAspThrLysValAsnPheTyrAlaTrpLysArgMetGluValGlyGlnGlnAla

12

50 60
 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu
 70 80
 5 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer
 90 100
 10 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer
 110 120
 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys
 130 140
 15 LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla
 150 160
 CysArgThrGlyAspArg
 166

20

wherein optionally 1-6 amino acids from the N-terminus and 1-5 from the C-terminus can be deleted from said EPO receptor agonists polypeptide;

25

wherein the N-terminus is joined to the C-terminus directly or through a linker capable of joining the N-terminus to the C-terminus and having new C- and N-termini at amino acids;

23-24	48-49	111-112
24-25	50-51	112-113
25-26	51-52	113-114
26-27	52-53	114-115
27-28	53-54	115-116
28-29	54-55	116-117
29-30	55-56	117-118
30-31	56-57	118-119
31-32	57-58	119-120
32-33	77-78	120-121
33-34	78-79	121-122
34-35	79-80	122-123
35-36	80-81	123-124
36-37	81-82	124-125
37-38	82-83	125-126
38-39	84-85	126-127
40-41	85-86	127-128
41-42	86-87	128-129
43-44	87-88	129-130
44-45	88-89	131-132
45-46	108-109	respectively; and
46-47	109-110	
47-48	110-111	

13

said EPO receptor agonist polypeptide may optionally be immediately preceded by (methionine⁻¹), (alanine⁻¹) or (methionine⁻², alanine⁻¹).

5

The more preferred breakpoints at which new C-terminus and N-terminus can be made are; 23-24, 24-25, 25-26, 27-28, 28-29, 29-30, 30-31, 31-32, 32-33, 33-34, 34-35, 35-36, 36-37, 37-38, 38-39, 40-41, 41-42, 42-43, 10 52-53, 53-54, 54-55, 55-56, 77-78, 78-79, 79-80, 80-81, 81-82, 82-83, 83-84, 84-85, 85-86, 86-87, 87-88, 88-89, 109-110, 110-111, 111-112, 112-113, 113-114, 114-115, 115-116, 116-117, 117-118, 118-119, 119-120, 120-121, 121-122, 122-123, 123-124, 124-125, 125-126, 126-127, 15 127-128, 128-129, 129-130, 130-131, and 131-132.

The most preferred breakpoints at which new C-terminus and N-terminus can be made are; 23-24, 24-25, 31-32, 32-33, 37-38, 38-39, 82-83, 83-84, 85-86, 86-87, 20 87-88, 125-126, 126-127, and 131-132.

The most preferred breakpoints include glycosylation sites, non-neutralizing antibodies, proteolyte cleavage sites.

25

The EPO receptor agonists of the present invention may contain amino acid substitutions, such as those disclosed in WO 94/24160 or one or more of the glycosylation sites at Asn²⁴, Asn⁸³, and Asn¹²⁶ are 30 changed to other amino acids such as but not limited to Asp or Glu, deletions and/or insertions. It is also intended that the EPO receptor agonists of the present invention may also have amino acid deletions at either/or both the N- and C- termini of the original 35 protein and or deletions from the new N- and/or C-termini of the sequence rearranged proteins in the formulas shown above.

14

A preferred embodiment of the present invention the linker (L) joining the N-terminus to the C-terminus is a polypeptide selected from the group consisting of:

- GlyGlyGlySer SEQ ID NO:123;
5 GlyGlyGlySerGlyGlyGlySer SEQ ID NO:124;
GlyGlyGlySerGlyGlyGlySerGlyGlyGlySer SEQ ID NO:
125;
SerGlyGlySerGlyGlySer SEQ ID NO:126;
GluPheGlyAsnMet SEQ ID NO:127;
10 GluPheGlyGlyAsnMet SEQ ID NO:128;
GluPheGlyGlyAsnGlyGlyAsnMet SEQ ID NO:129; and
GlyGlySerAspMetAlaGly SEQ ID NO:130.

- The present invention also encompasses recombinant
15 human EPO receptor agonists co-administered or
sequentially with one or more additional colony
stimulating factors (CSF) including, cytokines,
lymphokines, interleukins, hematopoietic growth factors
which include but are not limited to GM-CSF, G-CSF, c-
20 mpl ligand (also known as TPO or MGDF), M-CSF, IL-1, IL-
4, IL-2, IL-3, IL-5, IL 6, IL-7, IL-8, IL-9, IL-10, IL-
11, IL-12, IL-13, IL-15, LIF, human growth hormone, B-
cell growth factor, B-cell differentiation factor,
eosinophil differentiation factor and stem cell factor
25 (SCF) also known as steel factor or c-kit ligand (herein
collectively referred to as "factors"). These co-
administered mixtures may be characterized by having the
usual activity of both of the peptides or the mixture
may be further characterized by having a biological or
30 physiological activity greater than simply the additive
function of the presence of the EPO receptor agonists or
the second colony stimulating factor alone. The co-
administration may also provide an enhanced effect on
the activity or an activity different from that expected
35 by the presence of the EPO or the second colony
stimulating factor. The co-administration may also have
an improved activity profile which may include reduction

15

of undesirable biological activities associated with native human EPO. In addition to the list above, IL-3 variants taught in WO 94/12639 and WO 94/12638 fusion protein taught in WO 95/21197, and WO 95/21254 G-CSF
5 receptor agonists disclosed in WO 97/12977, c-mpl receptor agonists disclosed in WO 97/12978, IL-3 receptor agonists disclosed in WO 97/12979 and multi-functional receptor agonists taught in WO 97/12985 can be co-administered with the polypeptides of the present
10 invention. As used herein "IL-3 variants" refer to IL-3 variants taught in WO 94/12639 and WO 94/12638. As used herein "fusion proteins" refer to fusion protein taught in WO 95/21197, and WO 95/21254. As used herein "G-CSF receptor agonists" refer to G-CSF receptor agonists
15 disclosed in WO 97/12978. As used herein "c-mpl receptor agonists" refer to c-mpl receptor agonists disclosed in WO 97/12978. As used herein "IL-3 receptor agonists" refer to IL-3 receptor agonists disclosed in WO 97/12979. As used herein "multi-functional receptor
20 agonists" refer to multi-functional receptor agonists taught in WO 97/12985.

In addition, it is envisioned that in vitro uses would include the ability to stimulate bone marrow and
25 blood cell activation and growth before the expanded cells are infused into patients.

It is also envisioned that uses of EPO receptor agonists of the present invention would include blood
30 banking applications, where the EPO receptor agonists are given to a patient to increase the number of red blood cells and blood products removed from the patient, prior to some medical procedure, and the blood products stored and transfused back into the patient after the
35 medical procedure. Additionally, it is envisioned that uses of EPO receptor agonists would include giving the EPO receptor agonists to a blood donor prior to blood

16

donation to increase the number of red blood cells,
thereby allowing the donor to safely give more blood.

Brief Description of the Figures

Figure 1 schematically illustrates the sequence rearrangement of a protein. The N-terminus (N) and the C-terminus (C) of the native protein are joined through a linker, or joined directly. The protein is opened at a breakpoint creating a new N-terminus (new N) and a new C-terminus (new-C) resulting in a protein with a new linear amino acid sequence. A rearranged molecule may be synthesized *de novo* as linear molecule and not go through the steps of joining the original N-terminus and the C-terminus and opening of the protein at the breakpoint.

Figure 2 shows a schematic of Method I, for creating new proteins in which the original N-terminus and C-terminus of the native protein are joined with a linker and different N-terminus and C-terminus of the protein are created. In the example shown the sequence rearrangement results in a new gene encoding a protein with a new N-terminus created at amino acid 97 of the original protein, the original C-terminus (a.a. 174) joined to the amino acid 11 (a.a. 1- 10 are deleted) through a linker region and a new C-terminus created at amino acid 96 of the original sequence.

Figure 3 shows a schematic of Method II, for creating new proteins in which the original N-terminus and C-terminus of the native protein are joined without a linker and different N-terminus and C-terminus of the protein are created. In the example shown the sequence rearrangement results in a new gene encoding a protein with a new N-terminus created at amino acid 97 of the original protein, the original C-terminus (a.a. 174) joined to the original N-terminus and a new C-terminus created at amino acid 96 of the original sequence.

18

Figure 4 shows a schematic of Method III, for creating new proteins in which the original N-terminus and C-terminus of the native protein are joined with a linker and different N-terminus and C-terminus of the protein are created. In the example shown the sequence rearrangement results in a new gene encoding a protein with a new N-terminus created at amino acid 97 of the original protein, the original C-terminus (a.a. 174) joined to amino acid 1 through a linker region and a new C-terminus created at amino acid 96 of the original sequence.

Figure 5 shows a DNA sequence encoding human mature EPO based on the sequence of Lin et al. (PNAS 82:7580-7584, 1985).

Detailed Description of the Invention

Receptor agonists of the present invention may be useful in the treatment of diseases characterized by decreased levels of red blood cells of the hematopoietic system.

A EPO receptor agonist may be useful in the treatment or prevention of anemia. Many drugs may cause bone marrow suppression or hematopoietic deficiencies. Examples of such drugs are AZT, DDI, alkylating agents and anti-metabolites used in chemotherapy, antibiotics such as chloramphenicol, penicillin, gancyclovir, daunomycin and sulfa drugs, phenothiazones, tranquilizers such as meprobamate, analgesics such as aminopyrine and dipyrone, anti-convulsants such as phenytoin or carbamazepine, antithyroids such as propylthiouracil and methimazole and diuretics. EPO receptor agonists may be useful in preventing or treating the bone marrow suppression or hematopoietic deficiencies which often occur in patients treated with these drugs.

Hematopoietic deficiencies may also occur as a result of viral, microbial or parasitic infections and as a result of treatment for renal disease or renal failure, e.g., dialysis. The present peptide may be useful in treating such hematopoietic deficiency.

Another aspect of the present invention provides plasmid DNA vectors for use in the method of expression of these novel EPO receptor agonists. These vectors contain the novel DNA sequences described above which code for the novel polypeptides of the invention. Appropriate vectors which can transform host cells capable of expressing the EPO receptor agonists include expression vectors comprising nucleotide sequences coding for the EPO receptor agonists joined to transcriptional and translational regulatory sequences which are selected according to the host cells used.

20

Vectors incorporating modified sequences as described above are included in the present invention and are useful in the production of the modified EPO receptor agonist polypeptides. The vector employed in the method
5 also contains selected regulatory sequences in operative association with the DNA coding sequences of the invention and capable of directing the replication and expression thereof in selected host cells.

As another aspect of the present invention, there
10 is provided a method for producing the novel family of human EPO receptor agonists. The method of the present invention involves culturing suitable cells or cell line, which has been transformed with a vector containing a DNA sequence coding for expression of the
15 novel EPO receptor agonist polypeptide. Suitable cells or cell lines may include various strains of bacteria such as *E. coli*, yeast, mammalian cells, or insect cells may be utilized as host cells in the method of the present invention.

20

Other aspects of the present invention are methods and therapeutic compositions for treating the conditions referred to above. Such compositions comprise a therapeutically effective amount of one or more of the
25 EPO receptor agonists of the present invention in a mixture with a pharmaceutically acceptable carrier. This composition can be administered either parenterally, intravenously or subcutaneously. When administered, the therapeutic composition for use in
30 this invention is preferably in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such a parenterally acceptable protein solution, having due regard to pH, isotonicity, stability and the like, is within the skill of the art.

35

Administration will be in accordance with a dosage regimen that will be readily ascertained by the skilled,

2/

based on *in vivo* specific activity of the analog in comparison with human erythropoietin and based on what is now known in the art concerning the administration of human erythropoietin for inducing erythropoiesis and

5 treating various conditions, such as anemia, in humans, including anemia in patients suffering from renal failure. Dosage of an analog of the invention may vary somewhat from individual to individual, depending on the particular analog and its specific *in vivo* activity,

10 the route of administration, the medical condition, age, weight or sex of the patient, the patient's sensitivities to the analog or components of vehicle, and other factors which the attending physician will be capable of readily taking into account. With regard to

15 therapeutic uses of analogs of the invention, reference is made to U.S. Patent Nos. 4,703,008 and 4,835,260; see also the chapter on (recombinant) [des-Arg¹⁶⁶]human erythropoietin at pages 591-595 of the Physicians' Desk

20 Commercially available preparations of recombinant [des-Arg¹⁶⁶] human erythropoietin have 2,000, 3,000, 4,000 or 10,000 units of the glycohormone per mL in preservative-free aqueous solution with 2.5 mg/mL human serum albumin, 5.8 mg/mL sodium citrate, 5.8 mg/mL NaCl, and 0.06 mg/mL citric acid, pH 6.9 (+/-0.3).

25

Recombinantly produced EPO has proven especially useful for the treatment of patients suffering from impaired red blood cell production (Physicians Desk Reference (PDR). 1993 edition, pp 602-605). Recombinant

30 EPO has proven effective in treating anemia associated with chronic renal failure and HIV-Infected individuals suffering from lowered endogenous EPO levels related to therapy with Zidovudine (AZT) (See PDR, 1993 edition, at page 6002).

35

Modifications of the EPO protein which would improve its utility as a tool for diagnosis or treatment

22

of blood disorders are certainly desirable. In particular, modified forms of EPO exhibiting enhanced biological activity would be more effective and efficient than native EPO in the therapy setting when it is necessary to administer EPO to the patient, enabling administration less frequently and/or at a lower dose. Administration of reduced amounts of EPO would also presumably reduce the risk of adverse effects associated with EPO treatment, such as hypertension, seizures, headaches, etc. (See PDR, 1993 edition, at pp. 603-604). The EPO receptor agonists of the present invention may also have improved stability and hence increased half-life which would allow for the production of a non-glycosylated form of EPO in a bacterial expression system at a much lower cost. Due to its increased half-life this non-glycosylated form of EPO would have an increased in vivo activity compared to de-glycosylated EPO.

The therapeutic method and compositions may also include co-administration with other hematopoietic factors. A non-exclusive list of other appropriate hematopoietins, colony stimulating factors (CSFs) and interleukins for simultaneous or serial co-administration with the polypeptides of the present invention includes GM-CSF, G-CSF, c-mpl ligand (also known as TPO or MGDF), M-CSF, IL-1, IL-4, IL-2, IL-3, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-15, LIF, human growth hormone, B-cell growth factor, B-cell differentiation factor, eosinophil differentiation factor and stem cell factor (SCF) also known as steel factor or c-kit ligand (herein collectively referred to as "factors"), or combinations thereof. In addition to the list above, IL-3 variants taught in WO 94/12639 and WO 94/12638 fusion protein taught in WO 95/21197, and WO 95/21254 G-CSF receptor agonists disclosed in WO 97/12977, c-mpl receptor agonists disclosed in WO 97/12978, IL-3 receptor

23

agonists disclosed in WO 97/12979 and multi-functional receptor agonists taught in WO 97/12985 can be co-administered with the polypeptides of the present invention.

5

The EPO receptor agonists of the present invention may be useful in the mobilization of hematopoietic progenitors and stem cells in peripheral blood. Peripheral blood derived progenitors have been shown to
10 be effective in reconstituting patients in the setting of autologous marrow transplantation.

The EPO receptor agonists of the present invention may also be useful in the ex vivo expansion of
15 hematopoietic progenitors. Colony stimulating factors (CSFs), such as G-CSF, have been administered alone, co-administered with other CSFs, or in combination with bone marrow transplants subsequent to high dose chemotherapy to treat the anemia, neutropenia and
20 thrombocytopenia which are often the result of such treatment.

Another aspect of the invention provides methods of sustaining and/or expanding hematopoietic precursor cells which includes inoculating the cells into a
25 culture vessel which contains a culture medium that has been conditioned by exposure to a stromal cell line such as HS-5 (WO 96/02662, Roecklein and Torok-Strob, *Blood* 85:997-1105, 1995) that has been supplemented with a EPO receptor agonist of the present invention.

30

Determination of the Linker

35 The length of the amino acid sequence of the linker can be selected empirically or with guidance from structural information, or by using a combination of the two approaches.

24

When no structural information is available, a small series of linkers can be prepared for testing using a design whose length is varied in order to span a range from 0 to 50 Å and whose sequence is chosen in order to be consistent with surface exposure (hydrophilicity, Hopp & Woods, *Mol. Immunol.* **20**: 483-489, 1983; Kyte & Doolittle, *J. Mol. Biol.* **157**:105-132, 1982; solvent exposed surface area, Lee & Richards, *J. Mol. Biol.* **55**:379-400, 1971) and the ability to adopt the necessary conformation without deranging the configuration of the EPO receptor agonist (conformationally flexible; Karplus & Schulz, *Naturwissenschaften* **72**:212-213, (1985). Assuming an average of translation of 2.0 to 3.8 Å per residue, this would mean the length to test would be between 0 to 30 residues, with 0 to 15 residues being the preferred range. Exemplary of such an empirical series would be to construct linkers using a cassette sequence such as Gly-Gly-Gly-Ser repeated n times, where n is 1, 2, 3 or 4. Those skilled in the art will recognize that there are many such sequences that vary in length or composition that can serve as linkers with the primary consideration being that they be neither excessively long nor short (cf., Sandhu, *Critical Rev. Biotech.* **12**: 437-462, 1992); if they are too long, entropy effects will likely destabilize the three-dimensional fold, and may also make folding kinetically impractical, and if they are too short, they will likely destabilize the molecule because of torsional or steric strain.

30

Those skilled in the analysis of protein structural information will recognize that using the distance between the chain ends, defined as the distance between the c-alpha carbons, can be used to define the length of the sequence to be used, or at least to limit the number of possibilities that must be tested in an empirical selection of linkers. They will also recognize that it

35

25

is sometimes the case that the positions of the ends of the polypeptide chain are ill-defined in structural models derived from x-ray diffraction or nuclear magnetic resonance spectroscopy data, and that when
5 true, this situation will therefore need to be taken into account in order to properly estimate the length of the linker required. From those residues whose positions are well defined are selected two residues that are close in sequence to the chain ends, and the
10 distance between their c-alpha carbons is used to calculate an approximate length for a linker between them. Using the calculated length as a guide, linkers with a range of number of residues (calculated using 2 to 3.8Å per residue) are then selected. These linkers
15 may be composed of the original sequence, shortened or lengthened as necessary, and when lengthened the additional residues may be chosen to be flexible and hydrophilic as described above; or optionally the original sequence may be substituted for using a series
20 of linkers, one example being the "Gly-Gly-Gly-Ser" cassette approach mentioned above; or optionally a combination of the original sequence and new sequence having the appropriate total length may be used.

25

Determination of the Amino and Carboxyl Termini of EPO Receptor Agonists

Sequences of EPO receptor agonists capable of
30 folding to biologically active states can be prepared by appropriate selection of the beginning (amino terminus) and ending (carboxyl terminus) positions from within the original polypeptide chain while using the linker sequence as described above. Amino and carboxyl termini
35 are selected from within a common stretch of sequence, referred to as a breakpoint region, using the guidelines described below. A novel amino acid sequence is thus generated by selecting amino and carboxyl termini from

26

within the same breakpoint region. In many cases the selection of the new termini will be such that the original position of the carboxyl terminus immediately preceded that of the amino terminus. However, those skilled in the art will recognize that selections of termini anywhere within the region may function, and that these will effectively lead to either deletions or additions to the amino or carboxyl portions of the new sequence.

It is a central tenet of molecular biology that the primary amino acid sequence of a protein dictates folding to the three-dimensional structure necessary for expression of its biological function. Methods are known to those skilled in the art to obtain and interpret three-dimensional structural information using x-ray diffraction of single protein crystals or nuclear magnetic resonance spectroscopy of protein solutions. Examples of structural information that are relevant to the identification of breakpoint regions include the location and type of protein secondary structure (alpha and 3-10 helices, parallel and anti-parallel beta sheets, chain reversals and turns, and loops; Kabsch & Sander, *Biopolymers* **22**: 2577-2637, 1983; the degree of solvent exposure of amino acid residues, the extent and type of interactions of residues with one another (Chothia, *Ann. Rev. Biochem.* **53**:537-572; 1984) and the static and dynamic distribution of conformations along the polypeptide chain (Alber & Mathews, *Methods Enzymol.* **154**: 511-533, 1987). In some cases additional information is known about solvent exposure of residues; one example is a site of post-translational attachment of carbohydrate which is necessarily on the surface of the protein. When experimental structural information is not available, or is not feasible to obtain, methods are also available to analyze the primary amino acid sequence in order to make predictions of protein tertiary and secondary structure, solvent accessibility

27

and the occurrence of turns and loops. Biochemical methods are also sometimes applicable for empirically determining surface exposure when direct structural methods are not feasible; for example, using the
5 identification of sites of chain scission following limited proteolysis in order to infer surface exposure (Gentile & Salvatore, *Eur. J. Biochem.* **218**:603-621, 1993)

Thus using either the experimentally derived structural
10 information or predictive methods (e.g., Srinivisan & Rose *Proteins: Struct., Funct. & Genetics*, **22**: 81-99, 1995) the parental amino acid sequence is inspected to classify regions according to whether or not they are integral to the maintenance of secondary and tertiary
15 structure. The occurrence of sequences within regions that are known to be involved in periodic secondary structure (alpha and 3-10 helices, parallel and anti-parallel beta sheets) are regions that should be avoided. Similarly, regions of amino acid sequence that
20 are observed or predicted to have a low degree of solvent exposure are more likely to be part of the so-called hydrophobic core of the protein and should also be avoided for selection of amino and carboxyl termini. In contrast, those regions that are known or predicted
25 to be in surface turns or loops, and especially those regions that are known not to be required for biological activity, are the preferred sites for location of the extremes of the polypeptide chain. Continuous stretches of amino acid sequence that are preferred based on the
30 above criteria are referred to as a breakpoint region.

Materials and Methods

Recombinant DNA methods

35 Unless noted otherwise, all specialty chemicals were obtained from Sigma Co., (St. Louis, MO). Restriction endonucleases and T4 DNA ligase were

28

obtained from New England Biolabs (Beverly, MA) or
Boehringer Mannheim (Indianapolis, IN).

Transformation of *E. coli* strains

5

E. coli strains, such as DH5 α [™] (Life Technologies,
Gaithersburg, MD) and TG1 (Amersham Corp., Arlington
Heights, IL) are used for transformation of ligation
reactions and are the source of plasmid DNA for
10 transfecting mammalian cells. *E. coli* strains, such as
MON105 and JM101, can be used for expressing the EPO
receptor agonist of the present invention in the
cytoplasm or periplasmic space.

15 MON105 ATCC#55204: F⁻, lamda⁻, IN(rrnD, rrE)1, rpoD⁺,
rpoH358

DH5 α [™]: F⁻, phi80dlacZdeltaM15, delta(lacZYA-argF)U169,
deoR, recA1, endA1, hsdR17(rk⁻,mk⁺), phoA, supE44lamda⁻,
20 thi-1, gyrA96, relA1

TG1: delta(lac-pro), supE, thi-1, hsdD5/F'(traD36,
proA+B⁺, lacIq, lacZdeltaM15)

25 DH5 α [™] Subcloning efficiency cells are purchased as
competent cells and are ready for transformation using
the manufacturer's protocol, while both *E. coli* strains
TG1 and MON105 are rendered competent to take up DNA
using a CaCl₂ method. Typically, 20 to 50 mL of cells
30 are grown in LB medium (1% Bacto-tryptone, 0.5% Bacto-
yeast extract, 150 mM NaCl) to a density of
approximately 1.0 optical density unit at 600 nanometers
(OD₆₀₀) as measured by a Baush & Lomb Spectronic
spectrophotometer (Rochester, NY). The cells are
35 collected by centrifugation and resuspended in one-fifth
culture volume of CaCl₂ solution (50 mM CaCl₂, 10 mM
Tris-Cl, pH7.4) and are held at 4°C for 30 minutes. The

29

cells are again collected by centrifugation and resuspended in one-tenth culture volume of CaCl_2 solution. Ligated DNA is added to 0.2mL of these cells, and the samples are held at 4°C for 1 hour. The samples
5 are shifted to 42°C for two minutes and 1mL of LB is added prior to shaking the samples at 37°C for one hour. Cells from these samples are spread on plates (LB medium plus 1.5% Bacto-agar) containing either ampicillin (100 micrograms/mL, ug/mL) when selecting for ampicillin-
10 resistant transformants, or spectinomycin (75 ug/mL) when selecting for spectinomycin-resistant transformants. The plates are incubated overnight at 37°C. Single colonies are picked, grown in LB supplemented with appropriate antibiotic for 6-16 hours
15 at 37°C with shaking. Colonies are picked and inoculated into LB plus appropriate antibiotic (100 ug/mL ampicillin or 75 ug/mL spectinomycin) and are grown at 37°C while shaking. Before harvesting the cultures, 1 ul of cells are analyzed by PCR for the
20 presence of a EPO receptor agonist gene. The PCR is carried out using a combination of primers that anneal to the EPO receptor agonist gene and/or vector. After the PCR is complete, loading dye is added to the sample followed by electrophoresis as described earlier. A
25 gene has been ligated to the vector when a PCR product of the expected size is observed.

Methods for creation of genes with new N-terminus/C-terminus

30

Method I. Creation of genes with new N-terminus/C-terminus which contain a linker region.

Genes with new N-terminus/C-terminus which contain
35 a linker region separating the original C-terminus and N-terminus can be made essentially following the method described in L. S. Mullins, et al *J. Am. Chem. Soc.* **116**,

30

5529-5533 (1994). Multiple steps of polymerase chain reaction (PCR) amplifications are used to rearrange the DNA sequence encoding the primary amino acid sequence of the protein. The steps are illustrated in Figure 2.

5

In the first step, the primer set ("new start" and "linker start") is used to create and amplify, from the original gene sequence, the DNA fragment ("Fragment Start") that contains the sequence encoding the new N-terminal portion of the new protein followed by the linker that connects the C-terminal and N-terminal ends of the original protein. In the second step, the primer set ("new stop" and "linker stop") is used to create and amplify, from the original gene sequence, the DNA fragment ("Fragment Stop") that encodes the same linker as used above, followed by the new C-terminal portion of the new protein. The "new start" and "new stop" primers are designed to include the appropriate restriction enzyme recognition sites which allow cloning of the new gene into expression plasmids. Typical PCR conditions are one cycle 95°C melting for two minutes; 25 cycles 94°C denaturation for one minute, 50°C annealing for one minute and 72°C extension for one minute; plus one cycle 72°C extension for seven minutes. A Perkin Elmer GeneAmp PCR Core Reagents kit is used. A 100 ul reaction contains 100 pmole of each primer and one ug of template DNA; and 1x PCR buffer, 200 uM dGTP, 200 uM dATP, 200 uM dTTP, 200 uM dCTP, 2.5 units AmpliTaq DNA polymerase and 2 mM MgCl₂. PCR reactions are performed in a Model 480 DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT).

"Fragment Start" and "Fragment Stop", which have complementary sequence in the linker region and the coding sequence for the two amino acids on both sides of the linker, are joined together in a third PCR step to make the full-length gene encoding the new protein. The

3/

DNA fragments "Fragment Start" and "Fragment Stop" are resolved on a 1% TAE gel, stained with ethidium bromide and isolated using a Qiaex Gel Extraction kit (Qiagen). These fragments are combined in equimolar quantities, heated at 70°C for ten minutes and slow cooled to allow annealing through their shared sequence in "linker start" and "linker stop". In the third PCR step, primers "new start" and "new stop" are added to the annealed fragments to create and amplify the full-length new N-terminus/C-terminus gene. Typical PCR conditions are one cycle 95°C melting for two minutes; 25 cycles 94°C denaturation for one minute, 60°C annealing for one minute and 72°C extension for one minute; plus one cycle 72°C extension for seven minutes. A Perkin Elmer GeneAmp PCR Core Reagents kit is used. A 100 ul reaction contains 100 pmole of each primer and approximately 0.5 ug of DNA; and 1x PCR buffer, 200 uM dGTP, 200 uM dATP, 200 uM dTTP, 200 uM dCTP, 2.5 units AmpliTaq DNA polymerase and 2 mM MgCl₂. PCR reactions are purified using a Wizard PCR Preps kit (Promega).

Method II. Creation of genes with new N-terminus/C-terminus without a linker region.

New N-terminus/C-terminus genes without a linker joining the original N-terminus and C-terminus can be made using two steps of PCR amplification and a blunt end ligation. The steps are illustrated in Figure 3. In the first step, the primer set ("new start" and "P-bl start") is used to create and amplify, from the original gene sequence, the DNA fragment ("Fragment Start") that contains the sequence encoding the new N-terminal portion of the new protein. In the second step, the primer set ("new stop" and "P-bl stop") is used to create and amplify, from the original gene sequence, the DNA fragment ("Fragment Stop") that contains the sequence encoding the new C-terminal portion of the new

32

protein. The "new start" and "new stop" primers are designed to include appropriate restriction sites which allow cloning of the new gene into expression vectors. Typical PCR conditions are one cycle 95°C melting for two minutes; 25 cycles 94°C denaturation for one minute, 50°C annealing for 45 seconds and 72°C extension for 45 seconds. Deep Vent polymerase (New England Biolabs) is used to reduce the occurrence of overhangs in conditions recommended by the manufacturer. The "P-bl start" and "P-bl stop" primers are phosphorylated at the 5' end to aid in the subsequent blunt end ligation of "Fragment Start" and "Fragment Stop" to each other. A 100 ul reaction contained 150 pmole of each primer and one ug of template DNA; and 1x Vent buffer (New England Biolabs), 300 uM dGTP, 300 uM dATP, 300 uM dTTP, 300 uM dCTP, and 1 unit Deep Vent polymerase. PCR reactions are performed in a Model 480 DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT). PCR reaction products are purified using a Wizard PCR Preps kit (Promega).

20

The primers are designed to include appropriate restriction enzyme recognition sites which allow for the cloning of the new gene into expression vectors. Typically "Fragment Start" is designed to create a NcoI restriction site, and "Fragment Stop" is designed to create a HindIII restriction site. Restriction digest reactions are purified using a Magic DNA Clean-up System kit (Promega). Fragments Start and Stop are resolved on a 1% TAE gel, stained with ethidium bromide and isolated using a Qiaex Gel Extraction kit (Qiagen). These fragments are combined with and annealed to the ends of the ~ 3800 base pair NcoI/HindIII vector fragment of pMON3934 by heating at 50°C for ten minutes and allowed to slow cool. The three fragments are ligated together using T4 DNA ligase (Boehringer Mannheim). The result is a plasmid containing the full-length new N-terminus/C-terminus gene. A portion of the ligation reaction is

35

33

used to transform *E. coli* strain DH5a cells (Life Technologies, Gaithersburg, MD). Plasmid DNA is purified and sequence confirmed as below.

- 5 Method III. Creation of new N-terminus/C-terminus genes by tandem-duplication method

New N-terminus/C-terminus genes can be made based on the method described in R. A. Horlick, et al *Protein*
10 *Eng.* 5:427-431 (1992). Polymerase chain reaction (PCR) amplification of the new N-terminus/C-terminus genes is performed using a tandemly duplicated template DNA. The steps are illustrated in Figure 4.

15 The tandemly-duplicated template DNA is created by cloning and contains two copies of the gene separated by DNA sequence encoding a linker connecting the original C- and N-terminal ends of the two copies of the gene. Specific primer sets are used to create and amplify a
20 full-length new N terminus/C-terminus gene from the tandemly-duplicated template DNA. These primers are designed to include appropriate restriction sites which allow for the cloning of the new gene into expression vectors. Typical PCR conditions are one cycle 95°C
25 melting for two minutes; 25 cycles 94°C denaturation for one minute, 50°C annealing for one minute and 72°C extension for one minute; plus one cycle 72°C extension for seven minutes. A Perkin Elmer GeneAmp PCR Core Reagents kit (Perkin Elmer Corporation, Norwalk, CT) is
30 used. A 100 ul reaction contains 100 pmole of each primer and one ug of template DNA; and 1x PCR buffer, 200 uM dGTP, 200 uM dATP, 200 uM dTTP, 200 uM dCTP, 2.5 units AmpliTaq DNA polymerase and 2 mM MgCl₂. PCR reactions are performed in a Model 480 DNA thermal
35 cyclor (Perkin Elmer Corporation, Norwalk, CT). PCR reactions are purified using a Wizard PCR Preps kit (Promega).

34

DNA isolation and characterization

Plasmid DNA can be isolated by a number of
5 different methods and using commercially available kits
known to those skilled in the art. A few such methods
are shown herein. Plasmid DNA is isolated using the
Promega Wizard™ Miniprep kit (Madison, WI), the Qiagen
QIAwell Plasmid isolation kits (Chatsworth, CA) or
10 Qiagen Plasmid Midi kit. These kits follow the same
general procedure for plasmid DNA isolation. Briefly,
cells are pelleted by centrifugation (5000 x g), plasmid
DNA released with sequential NaOH/acid treatment, and
cellular debris is removed by centrifugation (10000 x
15 g). The supernatant (containing the plasmid DNA) is
loaded onto a column containing a DNA-binding resin, the
column is washed, and plasmid DNA eluted with TE. After
screening for the colonies with the plasmid of interest,
the *E. coli* cells are inoculated into 50-100 mLs of LB
20 plus appropriate antibiotic for overnight growth at 37°C
in an air incubator while shaking. The purified plasmid
DNA is used for DNA sequencing, further restriction
enzyme digestion, additional subcloning of DNA fragments
and transfection into mammalian, *E. coli* or other cells.

Sequence confirmation.

Purified plasmid DNA is resuspended in dH₂O and
quantitated by measuring the absorbance at 260/280 nm in
30 a Bausch and Lomb Spectronic 601 UV spectrometer. DNA
samples are sequenced using ABI PRISM™ DyeDeoxy™
terminator sequencing chemistry (Applied Biosystems
Division of Perkin Elmer Corporation, Lincoln City, CA)
kits (Part Number 401388 or 402078) according to the
35 manufacturers suggested protocol usually modified by the
addition of 5% DMSO to the sequencing mixture.
Sequencing reactions are performed in a Model 480 DNA
thermal cycler (Perkin Elmer Corporation, Norwalk, CT)

35

following the recommended amplification conditions. Samples are purified to remove excess dye terminators with Centri-Sep™ spin columns (Princeton Separations, Adelphia, NJ) and lyophilized. Fluorescent dye labeled
5 sequencing reactions are resuspended in deionized formamide, and sequenced on denaturing 4.75% polyacrylamide-8M urea gels using an ABI Model 373A automated DNA sequencer. Overlapping DNA sequence fragments are analyzed and assembled into master DNA
10 contigs using Sequencher v2.1 DNA analysis software (Gene Codes Corporation, Ann Arbor, MI).

Expression of EPO receptor agonists in mammalian cells

15 Mammalian Cell Transfection/Production of Conditioned Media

The BHK-21 cell line can be obtained from the ATCC (Rockville, MD). The cells are cultured in Dulbecco's
20 modified Eagle media (DMEM/high-glucose), supplemented to 2mM (mM) L-glutamine and 10% fetal bovine serum (FBS). This formulation is designated BHK growth media. Selective media is BHK growth media supplemented with 453 units/mL hygromycin B (Calbiochem, San Diego, CA).
25 The BHK-21 cell line was previously stably transfected with the HSV transactivating protein VP16, which transactivates the IE110 promoter found on the plasmid pMON3359 (See Hippenmeyer et al., *Bio/Technology*, pp.1037-1041, 1993). The VP16 protein drives expression
30 of genes inserted behind the IE110 promoter. BHK-21 cells expressing the transactivating protein VP16 are designated BHK-VP16. The plasmid pMON1118 (See Highkin et al., *Poultry Sci.*, **70**: 970-981, 1991) expresses the hygromycin resistance gene from the SV40 promoter. A
35 similar plasmid is available from ATCC, pSV2-hph.

BHK-VP16 cells are seeded into a 60 millimeter (mm) tissue culture dish at 3×10^5 cells per dish 24 hours

36

prior to transfection. Cells are transfected for 16 hours in 3 mL of "OPTIMEM"™ (Gibco-BRL, Gaithersburg, MD) containing 10 ug of plasmid DNA containing the gene of interest, 3 ug hygromycin resistance plasmid, 5 pMON1118, and 80 ug of Gibco-BRL "LIPOFECTAMINE"™ per dish. The media is subsequently aspirated and replaced with 3 mL of growth media. At 48 hours post-transfection, media from each dish is collected and assayed for activity (transient conditioned media). The 10 cells are removed from the dish by trypsin-EDTA, diluted 1:10 and transferred to 100 mm tissue culture dishes containing 10 mL of selective media. After approximately 7 days in selective media, resistant cells grow into colonies several millimeters in diameter. The colonies 15 are removed from the dish with filter paper (cut to approximately the same size as the colonies and soaked in trypsin/EDTA) and transferred to individual wells of a 24 well plate containing 1 mL of selective media. After the clones are grown to confluence, the 20 conditioned media is re-assayed, and positive clones are expanded into growth media.

Expression of EPO receptor agonists in *E. coli*

25 *E. coli* strain MON105 or JM101 harboring the plasmid of interest are grown at 37°C in M9 plus casamino acids medium with shaking in a air incubator Model G25 from New Brunswick Scientific (Edison, New Jersey). Growth is monitored at OD600 until it reaches 30 a value of 1, at which time nalidixic acid (10 milligrams/mL) in 0.1 N NaOH is added to a final concentration of 50 µg/mL. The cultures are then shaken at 37°C for three to four additional hours. A high degree of aeration is maintained throughout culture 35 period in order to achieve maximal production of the desired gene product. The cells are examined under a light microscope for the presence of inclusion bodies

37

(IB). One mL aliquots of the culture are removed for analysis of protein content by boiling the pelleted cells, treating them with reducing buffer and electrophoresis via SDS-PAGE (see Maniatis et al.

5 Molecular Cloning: A Laboratory Manual, 1982). The culture is centrifuged (5000 x g) to pellet the cells.

Additional strategies for achieving high-level expression of genes in *E. coli* can be found in Savvas,
10 C.M. (*Microbiological Reviews* 60;512-538, 1996).

Inclusion Body preparation, Extraction, Refolding,
Dialysis, DEAE Chromatography, and Characterization of
15 the EPO receptor agonists which accumulate as inclusion
bodies in *E. coli*.

Isolation of Inclusion Bodies:

20 The cell pellet from a 330 mL *E. coli* culture is resuspended in 15 mL of sonication buffer (10 mM 2-amino-2-(hydroxymethyl) 1,3-propanediol hydrochloride (Tris-HCl), pH 8.0 + 1 mM ethylenediaminetetraacetic acid (EDTA)). These resuspended cells are sonicated
25 using the microtip probe of a Sonicator Cell Disruptor (Model W-375, Heat Systems-Ultrasonics, Inc., Farmingdale, New York). Three rounds of sonication in sonication buffer followed by centrifugation are employed to disrupt the cells and wash the inclusion
30 bodies (IB). The first round of sonication is a 3 minute burst followed by a 1 minute burst, and the final two rounds of sonication are for 1 minute each.

Extraction and refolding of proteins from inclusion body
35 pellets:

38

Following the final centrifugation step, the IB pellet is resuspended in 10 mL of 50 mM Tris-HCl, pH 9.5, 8 M urea and 5 mM dithiothreitol (DTT) and stirred at room temperature for approximately 45 minutes to
5 allow for denaturation of the expressed protein.

The extraction solution is transferred to a beaker containing 70 mL of 5mM Tris-HCl, pH 9.5 and 2.3 M urea and gently stirred while exposed to air at 4°C for 18 to 48 hours to allow the proteins to refold. Refolding is
10 monitored by analysis on a Vydac (Hesperia, Ca.) C18 reversed phase high pressure liquid chromatography (RP-HPLC) column (0.46x25 cm). A linear gradient of 40% to 65% acetonitrile, containing 0.1% trifluoroacetic acid (TFA), is employed to monitor the refold. This gradient
15 is developed over 30 minutes at a flow rate of 1.5 mL per minute. Denatured proteins generally elute later in the gradient than the refolded proteins.

Purification:

20

Following the refold, contaminating *E. coli* proteins are removed by acid precipitation. The pH of the refold solution is titrated to between pH 5.0 and pH 5.2 using 15% (v/v) acetic acid (HOAc). This solution
25 is stirred at 4°C for 2 hours and then centrifuged for 20 minutes at 12,000 x g to pellet any insoluble protein.

The supernatant from the acid precipitation step is dialyzed using a Spectra/Por 3 membrane with a molecular
30 weight cut off (MWCO) of 3,500 daltons. The dialysis is against 2 changes of 4 liters (a 50-fold excess) of 10mM Tris-HCl, pH 8.0 for a total of 18 hours. Dialysis lowers the sample conductivity and removes urea prior to DEAE chromatography. The sample is then centrifuged (20
35 minutes at 12,000 x g) to pellet any insoluble protein following dialysis.

39

A Bio-Rad Bio-Scale DEAE2 column (7 x 52 mm) is used for ion exchange chromatography. The column is equilibrated in a buffer containing 10mM Tris-HCl, pH 8.0. The protein is eluted using a 0-to-500 mM sodium chloride (NaCl) gradient, in equilibration buffer, over 45 column volumes. A flow rate of 1 mL per minute is used throughout the run. Column fractions (2 mL per fraction) are collected across the gradient and analyzed by RP HPLC on a Vydac (Hesperia, Ca.) C18 column (0.46 x 25 cm). A linear gradient of 40% to 65% acetonitrile, containing 0.1% trifluoroacetic acid (TFA), is employed. This gradient is developed over 30 minutes at a flow rate of 1.5 mL per minute. Pooled fractions are then dialyzed against 2 changes of 4 liters (50-to-500-fold excess) of 10 mM ammonium acetate (NH₄Ac), pH 4.0 for a total of 18 hours. Dialysis is performed using a Spectra/Por 3 membrane with a MWCO of 3,500 daltons. Finally, the sample is sterile filtered using a 0.22µm syringe filter (µStar LB syringe filter, Costar, Cambridge, Ma.), and stored at 4°C.

In some cases the folded proteins can be affinity purified using affinity reagents such as mAbs or receptor subunits attached to a suitable matrix. Alternatively, (or in addition) purification can be accomplished using any of a variety of chromatographic methods such as: ion exchange, gel filtration or hydrophobic chromatography or reversed phase HPLC.

These and other protein purification methods are described in detail in Methods in Enzymology, Volume 182 'Guide to Protein Purification' edited by Murray Deutscher, Academic Press, San Diego, CA (1990).

Protein Characterization:

The purified protein is analyzed by RP-HPLC, electrospray mass spectrometry, and SDS-PAGE. The

40

protein quantitation is done by amino acid composition, RP-HPLC, and Bradford protein determination. In some cases tryptic peptide mapping is performed in conjunction with electrospray mass spectrometry to
5 confirm the identity of the protein.

Methylcellulose Assay

This assay reflects the ability of colony stimulating
10 factors to stimulate normal bone marrow cells to produce different types of hematopoietic colonies *in vitro* (Bradley et al., *Aust. Exp Biol. Sci.* **44**:287-300, 1966), Pluznik et al., *J. Cell Comp. Physio* **66**:319-324, 1965).

15

Methods

Approximately 30 mL of fresh, normal, healthy bone marrow aspirate are obtained from individuals following informed consent. Under sterile conditions samples are
20 diluted 1:5 with a 1X PBS (#14040.059 Life Technologies, Gaithersburg, MD.) solution in a 50 mL conical tube (#25339-50 Corning, Corning MD). Ficoll (Histopaque 1077 Sigma H-8889) is layered under the diluted sample and centrifuged, 300 x g for 30 min. The mononuclear
25 cell band is removed and washed two times in 1X PBS and once with 1% BSA PBS (CellPro Co., Bothel, WA). Mononuclear cells are counted and CD34+ cells are selected using the Ceprate LC (CD34) Kit (CellPro Co., Bothel, WA) column. This fractionation is performed
30 since all stem and progenitor cells within the bone marrow display CD34 surface antigen.

Cultures are set up in triplicate with a final volume of 1.0 mL in a 35 X 10 mm petri dish (Nunc#174926).

35 Culture medium is purchased from Terry Fox Labs. (HCC-4230 medium (Terry Fox Labs, Vancouver, B.C., Canada) and erythropoietin (Amgen, Thousand Oaks, CA.) is added

41

to the culture media. 3,000-10,000 CD34+ cells are added per dish. EPO receptor agonist proteins, in conditioned media from transfected mammalian cells or purified from conditioned media from transfected mammalian cells or *E. coli*, are added to give final concentrations ranging from .001 nM to 10 nM. Cultures are resuspended using a 3cc syringe and 1.0 mL is dispensed per dish. Control (baseline response) cultures received no colony stimulating factors. Positive control cultures received conditioned media (PHA stimulated human cells: Terry Fox Lab. H2400). Cultures are incubated at 37°C, 5% CO₂ in humidified air. Hematopoietic colonies which are defined as greater than 50 cells are counted on the day of peak response (days 10-11) using a Nikon inverted phase microscope with a 40x objective combination. Groups of cells containing fewer than 50 cells are referred to as clusters. Alternatively colonies can be identified by spreading the colonies on a slide and stained or they can be picked, resuspended and spun onto cytospin slides for staining.

Human Cord Blood Hematopoietic Growth Factor Assays

Bone marrow cells are traditionally used for in vitro assays of hematopoietic colony stimulating factor (CSF) activity. However, human bone marrow is not always available, and there is considerable variability between donors. Umbilical cord blood is comparable to bone marrow as a source of hematopoietic stem cells and progenitors (Broxmeyer et al., *PNAS USA* **89**:4109-113, 1992; Mayani et al., *Blood* **81**:3252-3258, 1993). In contrast to bone marrow, cord blood is more readily available on a regular basis. There is also a potential to reduce assay variability by pooling cells obtained fresh from several donors, or to create a bank of

42

cryopreserved cells for this purpose. By modifying the culture conditions, and/or analyzing for lineage specific markers, it is be possible to assay specifically for burst forming colonies (BFU-E) activity.

Methods

Mononuclear cells (MNC) are isolated from cord blood within 24 hr. of collection, using a standard density gradient (1.077 g/mL Histopaque). Cord blood MNC have been further enriched for stem cells and progenitors by several procedures, including immunomagnetic selection for CD14-, CD34+ cells; panning for SBA-, CD34+ fraction using coated flasks from Applied Immune Science (Santa Clara, CA); and CD34+ selection using a CellPro (Bothell, WA) avidin column. Either freshly isolated or cryopreserved CD34+ cell enriched fractions are used for the assay. Duplicate cultures for each serial dilution of sample (concentration range from 1 pM to 1204 pM) are prepared with 1x10⁴ cells in 1ml of 0.9% methylcellulose containing medium without additional growth factors (Methocult H4230 from Stem Cell Technologies, Vancouver, BC.). After culturing for 7-9 days, colonies containing >30 cells are counted.

Transfected cell lines:

Cell lines, such as BHK or the murine pro B cell line Baf/3, can be transfected with a colony stimulating factor receptor, such as the human EPO receptor which the cell line does not have. These transfected cell lines can be used to determine the cell proliferative activity and/or receptor binding.

EXAMPLE 1

Genes encoding the sequence rearranged EPO ligands can be constructed by any one of the methods described herein or by other recombinant methods known to those

43
skilled in the art. For the purpose of this example, the
site of permutation is between residues 131(Arg) and
132(Thr) of EPO. This is a site which is susceptible to
proteolytic cleavage, thereby indicating surface
5 exposure with a relatively high degree of flexibility.

In this example a new N-terminus and a new C-terminus is
created without a linker joining the original termini.
This is done, as described in Method II, in 2 steps of
10 PCR and a blunt end ligation.

In the first PCR step, using a vector containing the DNA
sequence of SEQ ID NO:120 as the template, and the
primers "new start" and "blunt start", a DNA fragment is
15 created which encodes the new N-terminus. This fragment
is termed "fragment start". The sequence underlined in
the new start primer is the NcoI restriction site.

New start primer = gcgcgcCCATGGACAATCACTGCTGAC SEQ ID
20 NO:131

Blunt start primer = TCTGTCCCCTGTCCT SEQ ID NO:132

In the second PCR step, using a vector containing the
DNA sequence of SEQ ID NO:120 as the template, and the
25 primers "new stop" and "blunt stop" create a DNA
fragment which encodes the new C-terminus. This
fragment is termed "fragment stop". The sequence
underlined in the new stop primer is the HindIII
restriction site.

30
New stop primer =
gcgcgcAAGCTTATTATCGGAGTGGAGCAGCTGAGGCCGCATC SEQ ID
NO:133

35 Blunt end primer = GCCCCACCAAGCCTCATCTGT SEQ ID NO:134

44

In the ligation step, the two fragments created in the two PCR reactions are ligated together, digested with NcoI and HindIII and cloned into an expression vector. The clones are screened by restriction analysis and DNA
5 sequenced to confirm the proper sequence. The primers can be designed to create restriction sites other than NcoI and HindIII to clone into other expression vectors.

10

EXAMPLE 2

The sequence rearranged EPO receptor agonists of the present invention can be assayed for bioactivity by the methods described herein or by other assays known to
15 those skilled in the art.

Additional techniques for the construction of the variant genes, recombinant protein expression, protein purification, protein characterization, biological
20 activity determination can be found in WO 94/12639, WO 94/12638, WO 95/20976, WO 95/21197, WO 95/20977, WO 95/21254 which are hereby incorporated by reference in their entirety.

25 All references, patents or applications cited herein are incorporated by reference in their entirety as if written herein.

Various other examples will be apparent to the
30 person skilled in the art after reading the present disclosure without departing from the spirit and scope of the invention. It is intended that all such other examples be included within the scope of the appended claims.

35

45

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: G. D. Searle and Company
- (ii) TITLE OF THE INVENTION: Novel Erythropoietin Receptor Agonists :
- (iii) NUMBER OF SEQUENCES: 134
- (iv) CORRESPONDENCE ADDRESS:
 (A) ADDRESSEE: G. D. Searle & Co.
 (B) STREET: P.O. Box 5110
 (C) CITY: Chicago
 (D) STATE: IL
 (E) COUNTRY: U. S. A.
 (F) ZIP: 60680
- (v) COMPUTER READABLE FORM:
 (A) MEDIUM TYPE: Diskette
 (B) COMPUTER: IBM Compatible
 (C) OPERATING SYSTEM: DOS
 (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 (A) APPLICATION NUMBER:
 (B) FILING DATE: 21-OCT-1997
 (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 (A) APPLICATION NUMBER: 60/034,044
 (B) FILING DATE: 25-OCT-1996
- (viii) ATTORNEY/AGENT INFORMATION:
 (A) NAME: Bennett, Dennis A
 (B) REGISTRATION NUMBER: 34,547
 (C) REFERENCE/DOCKET NUMBER: 2991/1
- (ix) TELECOMMUNICATION INFORMATION:
 (A) TELEPHONE: 314-737-6986
 (B) TELEFAX: 314-737-6972
 (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: None
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile
 1          5          10          15
Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu
 20          25          30
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser
 35          40          45
Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro
 50          55          60
Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg
 65          70          75          80
Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile
 85          90          95
Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala
100          105          110
Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly
115          120          125

```

```

Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly
130      135      140
Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu
145      150      155      160
Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu
      165      170

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr
1      5      10      15
Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val
20      25      30
Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Ser Glu
35      40      45
Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp
50      55      60
Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser
65      70      75      80
Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
85      90      95
Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp
100      105      110
Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys
115      120      125
Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly
130      135      140
Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg
145      150      155      160
Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn
165      170

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val
1      5      10      15
Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
20      25      30
Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
35      40      45
Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu
50      55      60
Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu
65      70      75      80
Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro
85      90      95
Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr
100      105      110
Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu
115      120      125
Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly
130      135      140
Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr
145      150      155      160
Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
165      170

```


47

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro
 1      5      10
Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln
 20     25     30
Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val
 35     40     45
Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro
 50     55     60
Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr
 65     70     75
Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro
 85     90     95
Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe
100    105    110
Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys
115    120    125
Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Ser
130    135    140
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
145    150    155    160
Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr
165    170

```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp
 1      5      10      15
Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln
 20     25     30
Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu
 35     40     45
Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
 50     55     60
Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr
 65     70     75     80
Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp
 85     90     95
Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg
100    105    110
Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu
115    120    125
Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala
130    135    140
Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu
145    150    155    160
Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr
165    170

```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

48

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr
 1      5      10      15
Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala
 20      25      30
Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg
 35      40      45
Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln
 50      55      60
Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu
 65      70      75      80
Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala
 85      90      95
Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys
100      105      110
Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr
115      120      125
Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro
130      135      140
Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu
145      150      155      160
Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly
165      170

```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys
 1      5      10      15
Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val
 20      25      30
Glu Val Trp Gln Gly Leu Ala Leu Ser Glu Ala Val Leu Arg Gly
 35      40      45
Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu
 50      55      60
His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
 65      70      75      80
Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala
 85      90      95
Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu
100      105      110
Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr
115      120      125
Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro
130      135      140
Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala
145      150      155      160
Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys
165      170

```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val
 1      5      10      15

```

SUBSTITUTE SHEET (rule 26)

```

Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu
  20      25
Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln
  35      40      45
Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His
  50      55      60
Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg
  65      70      75      80
Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser
  85      90      95
Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe
 100     105     110
Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly
 115     120     125
Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg
 130     135     140
Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys
 145     150     155     160
Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala
 165     170

```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn
  1      5      10      15
Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val
  20      25      30
Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala
  35      40      45
Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val
  50      55      60
Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala
  65      70      75      80
Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala
  85      90      95
Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg
 100     105     110
Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu
 115     120     125
Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu
 130     135     140
Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu
 145     150     155     160
Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu
 165     170

```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
  1      5      10      15
Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
  20      25      30
Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
  35      40      45
Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
  50      55      60
Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu

```

50

```

65      70      75      80
Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
85      90      95
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
100     105
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
115     120     125
Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile
130     135     140
Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala
145     150     155     160
Glu Asn Ile Thr Thr Gly Cys Ala Glu His
165     170

```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
1      5      10      15
Ala Trp Lys Arg Met Glu Val Gly Gln Ala Val Glu Val Trp Gln
20     25     30
Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu
35     40     45
Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys
50     55     60
Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Arg Ala Leu Gly
65     70     75     80
Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro
85     90     95
Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr
100    105    110
Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
115    120    125
Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys
130    135    140
Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu
145    150    155    160
Asn Ile Thr Thr Gly Cys Ala Glu His Cys
165    170

```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala
1      5      10      15
Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly
20     25     30
Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val
35     40     45
Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala
50     55     60
Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala
65     70     75     80
Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu
85     90     95
Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
100    105    110
Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg
115    120    125

```

51

```

Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp
130      135      140
Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn
145      150      155      160
Ile Thr Thr Gly Cys Ala Glu His Cys Ser
      165      170

```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp
1      5      10      15
Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu
20      25      30
Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn
35      40      45
Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val
50      55      60
Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln
65      70      75      80
Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg
85      90      95
Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn
100      105      110
Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr
115      120      125
Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser
130      135      140
Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
145      150      155      160
Thr Thr Gly Cys Ala Glu His Cys Ser Leu
      165      170

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys
1      5      10      15
Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
20      25      30
Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser
35      40      45
Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
50      55      60
Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys
65      70      75      80
Glu Ala Ile Ser Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr
85      90      95
Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe
100      105      110
Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
115      120      125
Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg
130      135      140
Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr
145      150      155      160
Thr Gly Cys Ala Glu His Cys Ser Leu Asn
      165      170

```

52

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg
 1           5           10           15
Met Glu Val Gly Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu
 20           25           30
Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
 35           40           45
Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly
 50           55           60
Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu
 65           70           75           80
Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile
 85           90           95
Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu
100          105          110
Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp
115          120          125
Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val
130          135          140
Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr
145          150          155          160
Gly Cys Ala Glu His Cys Ser Leu Asn Glu
165          170

```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met
 1           5           10           15
Glu Val Gly Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu
 20           25           30
Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln
 35           40           45
Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu
 50           55           60
Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala
 65           70           75           80
Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr
 85           90           95
Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg
100          105          110
Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg
115          120          125
Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu
130          135          140
Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly
145          150          155          160
Cys Ala Glu His Cys Ser Leu Asn Glu Asn
165          170

```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

53

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val
 1      5      10      15
Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Ser Glu
 20      25      30
Ala Val Leu Arg Gly Gln Ala Leu Val Asn Ser Ser Gln Pro Trp
 35      40      45
Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser
 50      55      60
Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
 65      70      75      80
Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp
 85      90      95
Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys
100      105      110
Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly
115      120      125
Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg
130      135      140
Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala
145      150      155      160
Glu His Cys Ser Leu Asn Glu Asn Ile Thr
165      170

```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
 1      5      10      15
Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
 20      25      30
Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu
 35      40      45
Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu
 50      55      60
Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro
 65      70      75      80
Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr
 85      90      95
Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu
100      105      110
Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly
115      120      125
Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr
130      135      140
Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu
145      150      155      160
His Cys Ser Leu Asn Glu Asn Ile Thr Val
165      170

```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln
 1      5      10      15

```

54

```

Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val
20 25 30
Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro
35 40 45
Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr
50 55 60
Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro
65 70 75 80
Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe
85 90 95
Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys
100 105 110
Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser
115 120 125
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
130 135 140
Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
145 150 155 160
Cys Ser Leu Asn Glu Asn Ile Thr Val Pro
165 170

```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
1 5 10 15
Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser
20 25 30
Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
35 40 45
Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys
50 55 60
Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr
65 70 75 80
Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe
85 90 95
Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
100 105 110
Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg
115 120 125
Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr
130 135 140
Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro
145 150 155 160
Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys
165 170

```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu
1 5 10 15
Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
20 25 30
Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly
35 40 45
Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu
50 55 60
Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile

```


65 70 55 75 80
 Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu
 85 90 95
 Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp
 100 105 110
 Arg Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val
 115 120 125
 Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr
 130 135 140
 Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp
 145 150 155 160
 Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg
 165 170

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu
 1 5 10 15
 Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln
 20 25 30
 Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu
 35 40 45
 Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala
 50 55 60
 Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr
 65 70 75 80
 Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg
 85 90 95
 Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg
 100 105 110
 Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu
 115 120 125
 Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly
 130 135 140
 Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr
 145 150 155 160
 Lys Val Asn Phe Tyr Ala Trp Lys Arg Met
 165 170

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser
 1 5 10 15
 Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro
 20 25 30
 Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg
 35 40 45
 Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile
 50 55 60
 Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala
 65 70 75 80
 Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly
 85 90 95
 Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly
 100 105 110
 Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu
 115 120 125

56

Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys
 130 135 140
 Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys
 145 150 155 160
 Val Asn Phe Tyr Ala Trp Lys Arg Met Glu
 165 170

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu
 1 5 10 15
 His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
 20 25 30
 Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala
 35 40 45
 Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu
 50 55 60
 Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr
 65 70 75 80
 Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Ser Ala Pro Pro
 85 90 95
 Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala
 100 105 110
 Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu
 115 120 125
 Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp
 130 135 140
 Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu
 145 150 155 160
 Ala Leu Leu Ser Glu Ala Val Leu Arg Gly
 165 170

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His
 1 5 10 15
 Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg
 20 25 30
 Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser
 35 40 45
 Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe
 50 55 60
 Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly
 65 70 75 80
 Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Ser Ala Pro Pro Arg
 85 90 95
 Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys
 100 105 110
 Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn
 115 120 125
 Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys
 130 135 140
 Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
 145 150 155 160
 Leu Leu Ser Glu Ala Val Leu Arg Gly Gln
 165 170

57

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val
 1          5          10          15
Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala
 20          25          30
Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala
 35          40          45
Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg
 50          55          60
Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu
 65          70          75          80
Ala Cys Arg Thr Gly Asp Arg Gly Gly Ser Ala Pro Pro Arg Leu
 85          90          95
Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu
100          105          110
Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu
115          120          125
Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg
130          135          140
Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu
145          150          155          160
Leu Ser Glu Ala Val Leu Arg Gly Gln Ala
165          170

```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
 1          5          10          15
Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
 20          25          30
Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
 35          40          45
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
 50          55          60
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
 65          70          75          80
Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile
 85          90          95
Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala
100          105          110
Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn
115          120          125
Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met
130          135          140
Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu
145          150          155          160
Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
165          170

```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

58

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```

Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys
 1          5          10          15
Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly
          20          25          30
Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro
          35          40          45
Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr
          50          55          60
Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
65          70          75          80
Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys
          85          90          95
Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu
          100          105          110
Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile
          115          120          125
Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu
          130          135          140
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser
145          150          155          160
Glu Ala Val Leu Arg Gly Gln Ala Leu Leu
          165          170

```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala
 1          5          10          15
Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala
          20          25          30
Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu
          35          40          45
Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
          50          55          60
Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg
65          70          75          80
Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp
          85          90          95
Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn
          100          105          110
Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr
          115          120          125
Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val
          130          135          140
Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu
145          150          155          160
Ala Val Leu Arg Gly Gln Ala Leu Leu Val
          165          170

```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```

Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val
 1          5          10          15

```

59

```

Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln
    20      25
Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg
    35      40      45
Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn
    50      55      60
Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr
    65      70      75      80
Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser
    85      90      95
Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
    100      105      110
Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val
    115      120      125
Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
    130      135      140
Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
    145      150      155      160
Val Leu Arg Gly Gln Ala Leu Leu Val Asn
    165      170

```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```

Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
  1      5      10
Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys
    20      25      30
Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr
    35      40      45
Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe
    50      55      60
Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
    65      70      75      80
Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg
    85      90      95
Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr
    100      105      110
Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro
    115      120      125
Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln
    130      135      140
Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val
    145      150      155      160
Leu Arg Gly Gln Ala Leu Leu Val Asn Ser
    165      170

```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```

Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly
  1      5      10      15
Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu
    20      25      30
Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile
    35      40      45
Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu
    50      55      60
Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp

```

60

65	70	75	80
Arg Gly Gly Gly Ser	Ala Pro Pro Arg	Leu Ile Cys Asp Ser	Arg Val
	85	90	95
Leu Glu Arg Tyr Leu	Leu Glu Ala Lys	Glu Ala Glu Asn Ile	Thr Thr
	100	105	110
Gly Cys Ala Glu His	Cys Ser Leu Asn Glu	Asn Ile Thr Val	Pro Asp
	115	120	125
Thr Lys Val Asn Phe Tyr	Ala Trp Lys Arg Met	Glu Val Gly Gln Gln	
	130	135	140
Ala Val Glu Val Trp Gln	Gly Leu Ala Leu Leu	Ser Glu Ala Val Leu	
	145	150	155
Arg Gly Gln Ala Leu	Leu Val Asn Ser Ser		
	165	170	

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Pro Trp Glu Pro Leu	Gln Leu His Val Asp	Lys Ala Val Ser Gly	Leu
1	5	10	15
Arg Ser Leu Thr Leu	Leu Arg Ala Leu	Gly Ala Gln Lys	Glu Ala
	20	25	30
Ile Ser Pro Pro Asp	Ala Ala Ser Ala Ala	Pro Leu Arg Thr	Ile Thr
	35	40	45
Ala Asp Thr Phe Arg	Lys Leu Phe Arg	Val Tyr Ser Asn	Phe Leu Arg
	50	55	60
Gly Lys Leu Lys Leu	Tyr Thr Gly Glu Ala	Cys Arg Thr Gly	Asp Arg
	65	70	75
Gly Gly Gly Ser Ala	Pro Pro Arg Leu Ile	Cys Asp Ser Arg	Val Leu
	85	90	95
Glu Arg Tyr Leu Leu	Glu Ala Lys Glu Ala	Glu Asn Ile Thr	Thr Gly
	100	105	110
Cys Ala Glu His Cys	Ser Leu Asn Glu Asn	Ile Thr Val Pro	Asp Thr
	115	120	125
Lys Val Asn Phe Tyr	Ala Trp Lys Arg Met	Glu Val Gly Gln	Gln Ala
	130	135	140
Val Glu Val Trp Gln	Gly Leu Ala Leu Leu	Ser Glu Ala Val	Leu Arg
	145	150	155
Gly Gln Ala Leu Leu	Val Asn Ser Ser Gln		
	165	170	

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Trp Glu Pro Leu Gln	Leu His Val Asp Lys	Ala Val Ser Gly	Leu Arg
1	5	10	15
Ser Leu Thr Thr Leu	Leu Arg Ala Leu	Gly Ala Gln Lys	Glu Ala Ile
	20	25	30
Ser Pro Pro Asp Ala	Ala Ala Ser Ala Ala	Pro Leu Arg Thr	Ile Thr Ala
	35	40	45
Asp Thr Phe Arg Lys	Leu Phe Arg Val Tyr	Ser Asn Phe Leu	Arg Gly
	50	55	60
Lys Leu Lys Leu Tyr	Thr Gly Glu Ala Cys	Arg Thr Gly Asp	Arg Gly
	65	70	75
Gly Gly Ser Ala Pro	Pro Arg Leu Ile Cys	Asp Ser Arg Val	Leu Glu
	85	90	95
Arg Tyr Leu Leu Glu	Ala Lys Glu Ala Glu	Asn Ile Thr Thr	Gly Cys
	100	105	110
Ala Glu His Cys Ser	Leu Asn Glu Asn	Ile Thr Val Pro	Asp Thr Lys
	115	120	125

61

```

Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val
 130      135
Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly
 145      150      155      160
Gln Ala Leu Leu Val Asn Ser Ser Gln Pro
      165      170

```

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```

Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser
 1      5      10      15
Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
 20      25      30
Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp
 35      40      45
Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys
 50      55      60
Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly
 65      70      75      80
Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg
 85      90      95
Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala
 100      105      110
Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val
 115      120      125
Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu
 130      135      140
Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln
 145      150      155      160
Ala Leu Leu Val Asn Ser Ser Gln Pro Trp
      165      170

```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```

Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala
 1      5      10      15
Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys
 20      25      30
Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr
 35      40      45
Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro
 50      55      60
Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu
 65      70      75      80
Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser
 85      90      95
Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala
 100      105      110
Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly
 115      120      125
Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val
 130      135      140
Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala
 145      150      155      160
Val Ser Gly Leu Arg Ser Leu Thr Thr Leu
      165      170

```

62

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```

Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala
 1      5      10
Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu
 20      25      30
Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr
 35      40      45
Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro
 50      55      60
Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala
 65      70      75      80
Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu
 85      90      95
Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp
100      105      110
Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu
115      120      125
Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn
130      135      140
Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val
145      150      155      160
Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
165      170

```

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```

Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser
 1      5      10
Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe
 20      25      30
Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly
 35      40      45
Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg
 50      55      60
Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys
 65      70      75      80
Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn
 85      90      95
Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys
100      105      110
Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
115      120      125
Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser
130      135      140
Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
145      150      155      160
Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg
165      170

```

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

6.3

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```

Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala
 1      5      10      15
Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg
 20      25      30
Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu
 35      40      45
Ala Cys Arg Thr Gly Asp Arg Gly Gly Ser Ala Pro Pro Arg Leu
 50      55      60
Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu
 65      70      75      80
Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu
 85      90      95
Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg
100      105      110
Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu
115      120      125
Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
130      135      140
Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly
145      150      155      160
Leu Arg Ser Leu Thr Leu Leu Arg Ala
165      170

```

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
 1      5      10      15
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
 20      25      30
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
 35      40      45
Cys Arg Thr Gly Asp Arg Gly Gly Ser Ala Pro Pro Arg Leu Ile
 50      55      60
Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala
 65      70      75      80
Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn
 85      90      95
Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met
100      105      110
Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu
115      120      125
Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln
130      135      140
Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu
145      150      155      160
Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
165      170

```

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```

Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro
 1      5      10      15

```

SUBSTITUTE SHEET (rule 26)

64

```

Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr
  20      25      30
Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
  35      40      45
Arg Thr Gly Asp Arg Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys
  50      55      60
Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu
  65      70      75      80
Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile
  85      90      95
Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu
 100     105     110
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser
 115     120     125
Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro
 130     135     140
Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg
 145     150     155     160
Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly
 165     170

```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```

Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu
  1      5      10      15
Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
  20      25      30
Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg
  35      40      45
Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp
  50      55      60
Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn
  65      70      75      80
Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr
  85      90      95
Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val
 100     105     110
Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu
 115     120     125
Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp
 130     135     140
Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser
 145     150     155     160
Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala
 165     170

```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg
  1      5      10      15
Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn
  20      25      30
Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr
  35      40      45
Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser
  50      55      60
Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile

```

65

```

65      70      75      80
Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val
85
Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
100
Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
115
Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu
130
Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu
145
Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln
165      170

```

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr
1      5      10      15
Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe
20      25      30
Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly
35      40      45
Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg
50      55      60
Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr
65      70      75      80
Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro
85      90      95
Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln
100      105      110
Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val
115      120      125
Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro
130      135      140
Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr
145      150      155      160
Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys
165      170

```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile
1      5      10      15
Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu
20      25      30
Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp
35      40      45
Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val
50      55      60
Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr
65      70      75      80
Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp
85      90      95
Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln
100      105      110
Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu
115      120      125

```

66

```

Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu
 130      135      140
Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr
 145      150      155      160
Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu
      165      170

```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```

Ile Ser Pro Pro Asp Ala Ala Ser Ala Pro Leu Arg Thr Ile Thr
 1      5      10      15
Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg
      20      25      30
Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg
      35      40      45
Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu
      50      55      60
Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly
      65      70      75      80
Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr
      85      90      95
Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala
      100      105      110
Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg
      115      120      125
Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln
      130      135      140
Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu
      145      150      155      160
Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala
      165      170

```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```

Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala
 1      5      10      15
Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly
      20      25      30
Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly
      35      40      45
Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu
      50      55      60
Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys
      65      70      75      80
Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys
      85      90      95
Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val
      100      105      110
Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly
      115      120      125
Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu
      130      135      140
His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu
      145      150      155      160
Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile
      165      170

```

67

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```

Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp
 1      5      10      15
Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys
 20      25      30
Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly
 35      40      45
Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg
 50      55      60
Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala
 65      70      75      80
Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val
 85      90      95
Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu
100      105      110
Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln
115      120      125
Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His
130      135      140
Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg
145      150      155      160
Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
165      170

```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```

Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr
 1      5      10      15
Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu
 20      25      30
Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly
 35      40      45
Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr
 50      55      60
Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu
 65      70      75      80
His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn
 85      90      95
Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val
100      105      110
Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala
115      120      125
Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val
130      135      140
Asp Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala
145      150      155      160
Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro
165      170

```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

68

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```

Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe
 1      5      10      15
Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys
 20      25      30
Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser
 35      40      45
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
 50      55      60
Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
 65      70      75      80
Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
 85      90      95
Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
100      105      110
Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
115      120      125
Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
130      135      140
Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
145      150      155      160
Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro
165      170

```

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```

Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg
 1      5      10      15
Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu
 20      25      30
Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala
 35      40      45
Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu
 50      55      60
Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys
 65      70      75      80
Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr
 85      90      95
Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln
100      105      110
Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu
115      120      125
Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys
130      135      140
Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly
145      150      155      160
Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp
165      170

```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys
 1      5      10      15

```

69

```

Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr
  20 25 30
Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro
  35 40 45
Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu
  50 55 60
Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser
  65 70 75 80
Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala
  85 90 95
Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly
 100 105 110
Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val
 115 120 125
Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala
 130 135 140
Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala
 145 150 155 160
Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala
 165 170

```

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

```

Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu
  1 5 10 15
Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr
  20 25 30
Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro
  35 40 45
Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala
  50 55 60
Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu
  65 70 75 80
Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp
  85 90 95
Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu
 100 105 110
Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn
 115 120 125
Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val
 130 135 140
Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln
 145 150 155 160
Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala
 165 170

```

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```

Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe
  1 5 10 15
Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly
  20 25 30
Glu Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg
  35 40 45
Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys
  50 55 60
Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn

```

```

65          70          75          80
Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys
      85          90          95
Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
      100         105         110
Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser
      115         120         125
Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser
      130         135         140
Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys
      145         150         155         160
Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser
      165         170

```

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```

Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg
 1          5          10          15
Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu
      20          25          30
Ala Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu
      35          40          45
Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu
      50          55          60
Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu
      65          70          75          80
Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg
      85          90          95
Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu
      100         105         110
Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
      115         120         125
Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly
      130         135         140
Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu
      145         150         155         160
Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala
      165         170

```

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
 1          5          10          15
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
      20          25          30
Cys Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile
      35          40          45
Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala
      50          55          60
Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn
      65          70          75          80
Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met
      85          90          95
Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu
      100         105         110
Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln
      115         120         125

```


Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu
 130 135 140
 Arg Ser Leu Thr Thr Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala
 145 150 155 160
 Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
 165 170

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 171 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr
 1 5 10 15
 Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
 20 25 30
 Arg Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys
 35 40 45
 Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu
 50 55 60
 Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile
 65 70 75 80
 Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu
 85 90 95
 Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser
 100 105 110
 Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro
 115 120 125
 Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg
 130 135 140
 Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Ala Lys Glu Ala
 145 150 155 160
 Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro
 165 170

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
 1 5 10 15
 Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg
 20 25 30
 Thr Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp
 35 40 45
 Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn
 50 55 60
 Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr
 65 70 75 80
 Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val
 85 90 95
 Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu
 100 105 110
 Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp
 115 120 125
 Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser
 130 135 140
 Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
 145 150 155 160
 Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu
 165 170

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 170 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn
 1           5           10           15
Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr
 20           25           30
Gly Asp Arg Gly Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser
 35           40           45
Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile
 50           55           60
Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val
 65           70           75           80
Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
 85           90           95
Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala
100           105           110
Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu
115           120           125
Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu
130           135           140
Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro
145           150           155           160
Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg
165           170

```

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

AATATCACGA CGGGCTGTGC TGAACACTGC AGCTTGAATG AGAATATCAC TGTCCCAGAC      60
ACCAAAGTTA ATTTCTATGC CTGGAAGAGG ATGGAGGTCG GGCAGCAGGC CGTAGAAGTC      120
TGGCAGGGCC TGGCCTGTCT GTCGGAAGCT GTCCTGCGGG GCCAGGCCCT GTTGGTCAAC      180
TCTTCCCAGC CGTGGGAGCC CCTGCAGCTG CATGTGGATA AAGCCGTCAG TGGCCTTCGC      240
AGCCTCACCA CTCTGCTTCG GGCTCTGGGA GCCCAGAAGG AAGCCATCTC CCCTCCAGAT      300
GCGGCTCAG CTGCTCCACT CCGAACAATC ACTGCTGACA CTTTCCGCAA ACTCTCCGA      360
GTCTACTCCA ATTTCTCTCG GGGAAAGCTG AAGCTGTACA CAGGGGAGGC CTGCAGGACA      420
GGGGACAGAT GAGGCGGCGG CTCCCCCAC CACGCCTCAT CTGTGACAGC CGAGTCCTGG      480
AGAGGTACCT CTTGGAGGCC AAGGAGGCC AG                                     512

```

(2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```

ATCACGACGG GCTGTGCTGA AACTGCAGC TTGAATGAGA ATATCACTGT CCCAGACACC      60
AAAGTTAATT TCTATGCCCTG GAAGAGGATG GAGGTCGGGC AGCAGGCCGT AGAAGTCTGG      120
CAGGGCCTGG CCTGTCTGTC GGAAGCTGTC CTGCGGGGCC AGGCCCTGTT GGTCAACTCT      180
TCCCAGCCGT GGGAGCCCTT GCAGCTGCAT GTGGATAAAG CCGTCAGTGG CTTTCGCAGC      240
CTCACCACTC TGCTTCGGGC TCTGGGAGCC CAGAAGGAAG CCATCTCCCC TCCAGATGCG      300
GCCTCAGCTG CTCCACTCCG AACAATCACT GCTGACACTT TCCGCAAACT CTTCCGAGTC      360
TACTCCAATT TCCTCCGGGG AAAGCTGAAG CTGTACACAG GGGAGGCCCT GAGGACAGGG      420
GACAGATGAG GCGGCGGCTC CCCCCACCAC GCCTCATCTG TGACAGCCGA GTCCTGGAGA      480
GGTACCTCTT GGAGGCCAAG GAGGCCGAGA AT                                     512

```

73

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

ACGACGGGCT	GTGCTGAACA	CTGCAGCTTG	AATGAGAATA	TCACTGTCCC	AGACACCAAA	60
GTTAATTTCT	ATGCCCTGGAA	GAGGATGGAG	GTCGGGCAGC	AGGCCGTAGA	AGTCTGGCAG	120
GGCCTGGCCC	TGCTGTCCGA	AGCTGTCCTG	CGGGGCCAGG	CCCTGTTGGT	CAACTCTTCC	180
CAGCCGTGGG	AGCCCTGCA	GCTGCATGTG	GATAAAGCCG	TCAGTGGCCT	TCGCAGCCTC	240
ACCACTCTGC	TTCGGGCTCT	GGGAGCCCAG	AAGGAAGCCA	TCTCCCCCTC	AGATGCGGCC	300
TCAGCTGCTC	CACTCCGAAC	AATCACTGCT	GACACTTTCC	GCAAACCTTT	CCGAGTCTAC	360
TCCAATTTC	TCCGGGGAAA	GCTGAAGCTG	TACACAGGGG	AGGCCTGCAG	GACAGGGGAC	420
AGATGAGGCG	GCGGCTCCCC	CCACCACGCC	TCATCTGTGA	CAGCCGAGTC	CTGGAGAGGT	480
ACCTCTTGGA	GGCCAAGGAG	GCCGAGAATA	TC			512

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

ACGGGCTGTG	CTGAACACTG	CAGCTTGAAT	GAGAATATCA	CTGTCCCAGA	CACCAAAGTT	60
AATTTCTATG	CCTGGAAGAG	GATGGAGGTC	GGGCAGCAGG	CCGTAGAAGT	CTGGCAGGGC	120
CTGGCCCTGC	TGTCGGAAGC	TGTCCTGCGG	GGCCAGGCC	TGTTGGTCAA	CTCTTCCAG	180
CCGTGGGAGC	CCCTGCAGCT	GCATGTGGAT	AAAGCCGTCA	GTGGCCTTCG	CAGCCTCACC	240
ACTCTGCTTC	GGGCTCTGGG	AGCCCAGAAG	GAAGCCATCT	CCCCCTCCAG	TGCGGCCTCA	300
GCTGCTCCAC	TCCGAACAAT	CACTGCTGAC	ACTTTCGCGA	AACTCTTCCG	AGTCTACTCC	360
AATTTCTCC	GGGAAAGCT	GAAGCTGTAC	ACAGGGGAGG	CCTGCAGGAC	AGGGGACAGA	420
TGAGGCGGCG	GCTCCCCCCA	CCACGCCTCA	TCTGTGACAG	CCGAGTCCTG	GAGAGGTACC	480
TCTTGGAGGC	CAAGGAGGCC	GAGAATATCA	CG			512

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GGCTGTGCTG	AACACTGCAG	CTTGAATGAG	AATATCACTG	TCCCAGACAC	CAAAGTTAAT	60
TTCTATGCCT	GGAAGAGGAT	GGAGGTCGGG	CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	120
GCCCTGTGCT	CGGAAGCTGT	CCTGCGGGGC	CAGGCCCTGT	TGGTCAATC	TTCCCAGCCG	180
TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	GCCGTCAGTG	GCCTTCGCAG	CCTCACCCT	240
CTGCTTCGGG	CTCTGGGAGC	CCAGAAGGAA	GCCATCTCCC	CTCCAGATGC	GGCCTCAGCT	300
GCTCCACTCC	GAACAATCAC	TGCTGACACT	TTCCGCAAAC	TCTTCCGAGT	CTACTCCAAT	360
TTCTTCCGGG	GAAAGCTGAA	GCTGTACACA	GGGGAGGCC	GCAGGACAGG	GGACAGATGA	420
GGCGGCGGCT	CCCCCACCA	CGCCTCATCT	GTGACAGCCG	AGTCTGGAG	AGGTACTCT	480
TGGAGGCCAA	GGAGGCCGAG	AATATCACGA	CG			512

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

TGTGCTGAAC	ACTGCAGCTT	GAATGAGAAT	ATCACTGTCC	CAGACACCAA	AGTTAATTTT	60
TATGCCTGGA	AGAGGATGGA	GGTCGGGCAG	CAGGCCGTAG	AAGTCTGGCA	GGGCCTGGCC	120
CTGCTGTCCG	AAGCTGTCCT	GCGGGGCCAG	GCCCTGTTGG	TCAACTCTTC	CCAGCCCTGG	180

74

GAGCCCCTGC	AGCTGCATGT	GGATAAAGCC	GTCAGTGGCC	TTCGCAGCCT	CACCACTCTG	240
CTTCGGGCTC	TGGGAGCCCA	GAAGGAAGCC	ATCTCCCCTC	CAGATGCGGC	CTCAGCTGCT	300
CCACTCCGAA	CAATCACTGC	TGACACTTTC	CGCAAACTCT	TCCGAGTCTA	CTCCAATTTT	360
CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	GAGGCCTGCA	GGACAGGGGA	CAGATGAGGC	420
GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	480
AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	GC			512

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GCTGAACACT	GCAGCTTGAA	TGAGAATATC	ACTGTCCCAG	ACACCAAAGT	TAATTTCTAT	60
GCCTGGAAGA	GGATGGAGGT	CGGGCAGCAG	GCCGTAGAAG	TCTGGCAGGG	CCTGGCCCTG	120
CTGTGCGAAG	CTGTCTGCG	GGGCCAGGCC	CTGTTGGTCA	ACTCTTCCCA	GCCGTGGGAG	180
CCCCTGCAAG	TGCATGTGGA	TAAAGCCGTC	AGTGGCCTTC	GCAGCCTCAC	CACTCTGCTT	240
CGGGCTCTGG	GAGCCAGAA	GGAAGCCATC	TCCCCTCCAG	ATGCGGCCTC	AGCTGCTCCA	300
CTCCGAACAA	TCACTGCTGA	CACTTTCCGC	AAACTCTTCC	GAGTCTACTC	CAATTTCTCT	360
CGGGGAAAGC	TGAAGCTGTA	CACAGGGGAG	GCCTGCAGGA	CAGGGGACAG	ATGAGGCGGC	420
GGCTCCCCCC	ACCACGCCTC	ATCTGTGACA	GCCGAGTCCT	GGAGAGGTAC	CTCTTGGAGG	480
CCAAGGAGGC	CGAGAATATC	ACGACGGGCT	GT			512

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GAACACTGCA	GCTTGAATGA	GAATATCACT	GTCCCAGACA	CCAAAGTTAA	TTTCTATGCC	60
TGGAAGAGGA	TGGAGGTCCG	GCAGCAGGCC	GTAGAAGTCT	GGCAGGGCCT	GGCCCTGCTG	120
TCGGAAGCTG	TCCTGCGGGG	CCAGGCCCTG	TTGGTCAACT	CTTCCAGGCC	GTGGGAGCCC	180
CTGCAGCTGC	ATGTGGATAA	AGCCGTCACT	GGCCTTCGCA	GCCTCACCAC	TCTGCTTCGG	240
GCTCTGGGAG	CCCAGAAGGA	AGCCATCTCC	CCTCCAGATG	CGGCCTCAGC	TGCTCCACTC	300
CGAACAATCA	CTGCTGACAC	TTTCCGCAAA	CTCTTCCGAG	TCTACTCCAA	TTTCTCCCGG	360
GGAAAGCTGA	AGCTGTACAC	AGGGGAGGCC	TGCAGGACAG	GGGACAGATG	AGGCGGCGGC	420
TCCCCCACCC	ACGCCTCATC	TGTGACAGCC	GAGTCCTGGA	GAGGTACCTC	TTGGAGGCCA	480
AGGAGGCCGA	GAATATCAGC	ACGGCTGTG	CT			512

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CACTGCAGCT	TGAATGAGAA	TATCACTGTC	CCAGACACCA	AAGTTAATTT	CTATGCCTGG	60
AAGAGGATGG	AGGTGCGGCA	GCAGGCCGTA	GAAGTCTGGC	AGGGCCTGGC	CCTGCTGTCT	120
GAAGCTGTCC	TGCGGGGCCA	GGCCCTGTTG	GTCAACTCTT	CCCAGCCGTG	GGAGCCCCCTG	180
CAGCTGCATG	TGGATAAAGC	CGTCAGTGGC	CTTCGCAAGC	TCACCACTCT	GCTTCGGGCT	240
CTGGGAGCCC	AGAAGGAAGC	CATCTCCCTT	CCAGATGCGG	CCTCAGCTGC	TCCACTCCGA	300
ACAATCACTG	CTGACACTTT	CCGCAAACTC	TTCCGAGTCT	ACTCCAATTT	CCTCCGGGGA	360
AAGCTGAAGC	TGTACACAGG	GGAGGCTGTC	AGGACAGGGG	ACAGATGAGG	CGGCGGCTCC	420
CCCCACCACG	CCTCATCTGT	GACAGCCGAG	TCCTGGAGAG	GTACCTCTTG	GAGGCCAAGG	480
AGGCCGAGAA	TATCACGACG	GGCTGTGCTG	AA			512

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

SUBSTITUTE SHEET (rule 26)

75

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TGCAGCTTGA	ATGAGAATAT	CACTGTCCCA	GACACCAAAG	TTAATTCTTA	TGCCTGGAAG	60
AGGATGGAGG	TCGGGCAGCA	GGCCGTAGAA	GTCTGGCAGG	GCCTGGCCCT	GCTGTGGAA	120
GCTGTCTTGC	GGGGCCAGGC	CCTGTTGGTC	AACTCTTCCC	AGCCGTGGGA	GGCCCTGCAG	180
CTGCATGTGG	ATAAAGCCGT	CAGTGGCCTT	CGCAGCCTCA	CCACTCTGCT	TCGGGCTCTG	240
GGAGCCCAGA	AGGAAGCCAT	CTCCCCTCCA	GATGCGGCCT	CAGCTGCTCC	ACTCCGAACA	300
ATCACTGCTG	ACACTTTCCG	CAAACCTCTC	CGAGTCTACT	CCAATTCTCT	CCGGGGAAAG	360
CTGAAGCTGT	ACACAGGGGA	GGCCTGCAGG	ACAGGGGACA	GATGAGGCGG	CGGCTCCCCC	420
CACCACGCCCT	CATCTGTGAC	AGCCGAGTCC	TGGAGAGGTA	CCTCTTGGAG	GCCAAGGAGG	480
CCGAGAATAT	CACGACGGGC	TGTGCTGAAC	AC			512

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

AGCTTGAATG	AGAATATCAC	TGTCCCAGAC	ACCAAAGTTA	ATTTCTATGC	CTGGAAGAGG	60
ATGGAGGTCG	GGCAGCAGGC	CGTAGAAGTC	TGGCAGGGCC	TGGCCCTGCT	GTCGGAAGCT	120
GTCCTGCGGG	GCCAGGCCCT	GTTGGTCAAC	TCTTCCCAGC	CGTGGGAGCC	CCTGCAGCTG	180
CATGTGGATA	AAGCCGTCAG	TGGCCTTCGC	AGCCTCACCA	CTCTGCTTCG	GGCTCTGGGA	240
GCCAGAAAGG	AAGCCATCTC	CCCTCCAGAT	GCGGCCTCAG	CTGCTCCACT	CCGAACAATC	300
ACTGCTGACA	CTTCCGCAA	ACTCTTCCGA	GTCTACTCCA	ATTTCTCCCG	GGGAAAGCTG	360
AAGCTGTACA	CAGGGGAGGC	CTGCAGGACA	GGGGACAGAT	GAGGCGGCGG	CTCCCCCAC	420
CACGCCCTCAT	CTGTGACAGC	CGAGTCCTGG	AGAGGTACCT	CTTGGAGGCC	AAGGAGGCCG	480
AGAATATCAC	GACGGGCTGT	GCTGAACACT	GC			512

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

TTGAATGAGA	ATATCACTGT	CCCAGACACC	AAAGTTAATT	TCTATGCCTG	GAAGAGGATG	60
GAGGTGCGGC	AGCAGGCCCT	AGAAGTCTGG	CAGGGCCTGG	CCCTGCTGTC	GGAAGCTGTC	120
CTGCGGGGCC	AGGCCCTGTT	GGTCAACTCT	TCCCAGCCGT	GGGAGCCCTT	GCAGCTGCAT	180
GTGGATAAAG	CCGTCAGTGG	CCTTCGCAGC	CTCACCCTC	TGCTTCGGGC	TCTGGGAGCC	240
CAGAAGGAAG	CCATCTCCCC	TCCAGATGCG	GCCTCAGCTG	CTCCACTCCG	AACAATCACT	300
GCTGACACTT	TCCGCAAAC	CTTCCGAGTC	TACTCCAATT	TCCTCCGGGG	AAAGCTGAAG	360
CTGTACACAG	GGGAGGCCCT	CAGGACAGGG	GACAGATGAG	GCGGCGGCTC	CCCCACCAC	420
GCCTCATCTG	TGACAGCCGA	GTCTTGGAGA	GGTACCTCTT	GGAGGCCAAG	GAGGCCGAGA	480
ATATCACGAC	GGGCTGTGCT	GAACACTGCA	GC			512

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

AATGAGAATA	TCAGTGTCCC	AGACACCAAA	GTAAATTCTT	ATGCCTGGAA	GAGGATGGAG	60
GTGCGGCAGC	AGGCCGTAGA	AGTCTGGCAG	GGCCTGGCCC	TGCTGTGCGA	AGCTGTCCTG	120
CGGGGCCAGG	CCCTGTTGGT	CAACTCTTCC	CAGCCGTGGG	AGCCCTTGCA	GCTGCATGTG	180
GATAAAGCCG	TCAGTGGCCT	TCGCAGCCTC	ACCACTCTGC	TTCGGGCTCT	GGGAGCCCAG	240
AAGGAAGCCA	TCTCCCCTCC	AGATGCGGCC	TCAGCTGCTC	CACTCCGAAC	AATCACTGCT	300
GACACTTTCC	GCAAACCTCT	CCGAGTCTAC	TCCAATTTC	TCCGGGGAAA	GCTGAAGCTG	360
TACACAGGGG	AGGCCTGCAG	GACAGGGGAC	AGATGAGGCG	GCGGCTCCCC	CCACCACGCC	420
TCATCTGTGA	CAGCCGAGTC	CTGGAGAGGT	ACCTCTTGGA	GGCCAAGGAG	GCCGAGAATA	480
TCACGACGGG	CTGTGCTGAA	CACTGCAGCT	TG			512

SUBSTITUTE SHEET (rule 26)

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GAGAATATCA	CTGTCCCAGA	CACCAAAGTT	AATTTCTATG	CCTGGAAGAG	GATGGAGGT	60
GGGCAGCAGG	CCGTAGAAGT	CTGGCAGGGC	CTGGCCCTGC	TGTCGGAAGC	TGTCCTGCGG	120
GGCCAGGCC	TGTTGGTCAA	CTCTTCCCAG	CCGTGGGAGC	CCCTGCAGCT	GCATGTGGAT	180
AAAGCCGTCA	GTGGCCTTCG	CAGCCTCACC	ACTCTGCTTC	GGGCTCTGGG	AGCCCAGAAG	240
GAAGCCATCT	CCCTCCAGA	TGCGGCCTCA	GCTGCTCCAC	TCCGAACAAT	CACTGCTGAC	300
ACTTTCCGCA	AACTCTTCCG	AGTCTACTCC	AATTTCTCTC	GGGGAAAGCT	GAAGCTGTAC	360
ACAGGGGAGG	CCTGCAGGAC	AGGGGACAGA	TGAGGCGGCG	GCTCCCCCA	CCACGCCTCA	420
TCTGTGACAG	CCGAGTCCTG	GAGAGGTACC	TCTTGGAGGC	CAAGGAGGCC	GAGAATATCA	480
CGACGGGCTG	TGCTGAACAC	TGCAGCTTGA	AT			512

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

AATATCACTG	TCCCAGACAC	CAAAGTTAAT	TTCTATGCCT	GGAAGAGGAT	GGAGGTCGGG	60
CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	GCCCTGCTGT	CGGAAGCTGT	CCTGCGGGGC	120
CAGGCCCTGT	TGGTCAATC	TTCCCAGCCG	TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	180
GCCGTCAGTG	GCCTTCGCAG	CCTCACCCT	CTGCTTCGGG	CTCTGGGAGC	CCAGAAGGAA	240
GCCATCTCCC	CTCCAGATGC	GGCTCAGCT	GCTCCACTCC	GAACAATCAC	TGCTGACACT	300
TTCCGCAAC	TCTTCCGAGT	CTACTCCAAT	TTCTCCGGG	GAAAGCTGAA	GCTGTACACA	360
GGGGAGGCC	GCAGGACAGG	GGACAGATGA	GGCGGCGGCT	CCCCCACC	CGCCTCATCT	420
GTGACAGCCG	AGTCCTGGAG	AGGTACCTCT	TGGAGGCCAA	GGAGGCCGAG	AATATCACGA	480
CGGGCTGTGC	TGAACACTGC	AGCTTGAATG	AG			512

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

ATCACTGTCC	CAGACACCAA	AGTTAATTTC	TATGCCTGGA	AGAGGATGGA	GGTCGGGCAG	60
CAGGCCGTAG	AAGTCTGGCA	GGGCCTGGCC	CTGCTGTCCG	AAGCTGTCCT	GCGGGGCCAG	120
GCCCTGTTGG	TCAACTCTTC	CCAGCCGTGG	GAGCCCTGTC	AGCTGCATGT	GGATAAAGCC	180
GTCAGTGGCC	TTCCGAGCCT	CACCACTCTG	CTTCGGGCTC	TGGGAGCCCA	GAAGGAAGCC	240
ATCTCCCCTC	CAGATGCGGC	CTCAGCTGCT	CCACTCCGAA	CAATCACTGC	TGACACTTTC	300
CGCAAACCTC	TCCGAGTCTA	CTCCAATTTC	CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	360
GAGGCCTGCA	GGACAGGGGA	CAGATGAGGC	GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	420
ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	480
GCTGTGCTGA	ACACTGCAGC	TTGAATGAGA	AT			512

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

ACTGTCCCAG	ACACCAAAGT	TAATTTCTAT	GCCTGGAAGA	GGATGGAGGT	CGGGCAGCAG	60
GCCGTAGAAG	TCTGGCAGGG	CCTGGCCCTG	CTGTCCGAAG	CTGTCCTGCG	GGGCCAGGCC	120
CTGTTGGTCA	ACTCTTCCCA	GCCGTGGGAG	CCCCTGCAGC	TGCATGTGGA	TAAAGCCGTG	180

77

AGTGGCCTTC	GCAGCCTCAC	CACTCTGCTT	CGGGCTCTGG	GAGCCCAGAA	GGAAGCCATC	240
TCCCTCCAG	ATGCGGCCTC	AGCTGCTCCA	CTCCGAACAA	TCCTGCTGA	CACTTTCCGC	300
AAACTCTTCC	GAGTCTACTC	CAATTTCTCT	CGGGGAAAGC	TGAAGCTGTA	CACAGGGGAG	360
GCCTGCAGGA	CAGGGACAG	ATGAGGCGGC	GGCTCCCCC	ACCACGCCTC	ATCTGTGACA	420
GCCGAGTCTC	GGAGAGGTAC	CTCTTGGAGG	CCAAGGAGGC	CGAGAATATC	ACGACGGGCT	480
GTGCTGAACA	CTGCAGCTTG	AATGAGAATA	ATC			513

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GTCCAGACA	CCAAAGTTAA	TTTCTATGCC	TGGAAGAGGA	TGGAGGTCGG	GCAGCAGGCC	60
GTAGAAGTCT	GGCAGGGCCT	GGCCCTGCTG	TCGGAAGCTG	TCCTGCGGGG	CCAGGCCCTG	120
TTGGTCAACT	CTTCCCAGCC	GTGGGAGCCC	CTGCAGCTGC	ATGTGGATAA	AGCCGTCAGT	180
GGCCTTCGCA	GCCTCACCAC	TCTGCTTCGG	GCTCTGGGAG	CCCAGAAGGA	AGCCATCTCC	240
CCTCCAGATG	CGGCCTCAGC	TGCTCCACTC	CGAACAATCA	CTGCTGACAC	TTTCCGCAAA	300
CTCTTCCGAG	TCTACTCCAA	TTTCTCCGG	GGAAAGCTGA	AGCTGTACAC	AGGGGAGGCC	360
TGCAGGACAG	GGGACAGATG	AGGCGGCGGC	TCCCCCACC	ACGCCTCATC	TGTGACAGCC	420
GAGTCCTGGA	GAGGTACCTC	TGGAGGCCA	AGGAGGCCGA	GAATATCAG	ACGGGCTGTG	480
CTGAACACTG	CAGCTGAAT	GAGAATAATC	ACT			513

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CCAGACACCA	AAGTTAATTT	CTATGCCTGG	AAGAGGATGG	AGGTCGGGCA	GCAGGCCGTA	60
GAAGTCTGGC	AGGGCCTGGC	CCTGCTGTGG	GAAGCTGTCC	TGCGGGGCCA	GGCCCTGTTG	120
GTCAACTCTT	CCCAGCCGTG	GGAGCCCCTG	CAGCTGCATG	TGGATAAAGC	CGTCAGTGCC	180
CTTCGCAGCC	TCACCACTCT	GCTTCGGGCT	CTGGGAGCCC	AGAAGGAAGC	CATCTCCCTT	240
CCAGATGCGG	CCTCAGCTGC	TCCACTCCGA	ACAATCACTG	CTGACACTTT	CCGCAAACTC	300
TTCCGAGTCT	ACTCCAATTT	CCTCCGGGGA	AAGCTGAAGC	TGTACACAGG	GGAGGCCTGC	360
AGGACAGGGG	ACAGATGAGG	CGGCGGCTCC	CCCCACCACG	CCTCATCTGT	GACAGCCGAG	420
TCCTGGAGAG	GTACCTCTTG	GAGGCCAAGG	AGGCCGAGAA	TATCACGACG	GGCTGTGCTG	480
AACACTGCAG	CTTGAATGAG	AATAATCACT	GTC			513

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GACACCAAAG	TTAATTCTTA	TGCCTGGAAG	AGGATGGAGG	TCGGGCAGCA	GGCCGTAGAA	60
GTCTGGCAGG	GCCTGGCCCT	GCTGTGCGAA	GCTGTCTCTG	GGGGCCAGGC	CCTGTTGGTC	120
AACTCTTCCC	AGCCGTGGGA	GCCCCGTCAG	CTGCATGTGG	ATAAAGCCGT	CAGTGGCCTT	180
CGCAGCCTCA	CCACTCTGCT	TCGGGCTCTG	GGAGCCCAGA	AGGAAGCCAT	CTCCCCCTCCA	240
GATCGGCCTT	CAGCTGCTCC	ACTCCGAACA	ATCACTGCTG	ACACTTTCCG	CAAACTCTTC	300
CGAGTCTACT	CCAATTTCCT	CCGGGGAAAG	CTGAAGCTGT	ACACAGGGGA	GGCCTGCAGG	360
ACAGGGGACA	GATGAGGCGG	CGGCTCCCCC	CACCACGCCT	CATCTGTGAC	AGCCGAGTCC	420
TGGAGAGGTA	CCTCTTGGAG	GCCAAGGAGG	CCGAGAATAT	CACGACGGGC	TGTGCTGAAC	480
ACTGCAGCTT	GAATGAGAAT	AATCACTGTC	CCA			513

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

78

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

AGGATGGAGG	TCGGGCAGCA	GGCCGTAGAA	GTCTGGCAGG	GCCTGGCCCT	GCTGTCGGAA	60
GCTGTCTGTC	GGGGCCAGGC	CCTGTTGGTC	AACTCTTCCC	AGCCGTGGGA	GCCCCTGCAG	120
CTGCATGTGG	ATAAAGCCGT	CAGTGGCCTT	CGCAGCCTCA	CCACTCTGCT	TCGGGCTCTG	180
GGAGCCCAGA	AGGAAGCCAT	CTCCCCCTCA	GATGCGGCCT	CAGCTGCTCC	ACTCCGAACA	240
ATCACTGCTG	ACACTTTCCG	CAAACCTCTC	CGAGTCTACT	CCAATTTCCT	CCGGGGAAAG	300
CTGAAGCTGT	ACACAGGGGA	GGCCTGCAGG	ACAGGGGACA	GATGAGGCGG	CGGCTCCCCC	360
CACCACGCCCT	CATCTGTGAC	AGCCGAGTCC	TGGAGAGGTA	CCTCTTGGAG	GCCAAGGAGG	420
CCGAGAATAT	CACGACGGGC	TGTGCTGAAC	ACTGCAGCTT	GAATGAGAAT	AATCACTGTC	480
CCAGACACCA	AAGTTAATTT	CTATGCCTGG	AAG			513

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

ATGGAGGTCG	GGCAGCAGGC	CGTAGAAGTC	TGGCAGGGCC	TGGCCCTGCT	GTCGGAAGCT	60
GTCTTGCGGG	GCCAGGCCCT	GTGGTCAAC	TCTTCCCAGC	CGTGGGAGCC	CCTGCAGCTG	120
CATGTGGATA	AAGCCGTACG	TGGCCTTCGC	AGCCTCACCA	CTCTGCTTCG	GGCTCTGGGA	180
GCCCAGAAGG	AAGCCATCTC	CCCTCCAGAT	GCGGCCTCAG	CTGTCCACT	CCGAACAATC	240
ACTGTGACA	CTTTCCGCAA	ACTCTTCCGA	GTCTACTCCA	ATTTCTCCG	GGGAAAGCTG	300
AAGCTGTACA	CAGGGGAGGC	CTGCAGGACA	GGGGACAGAT	GAGGCGGCGG	CTCCCCCAC	360
CACGCCTCAT	CTGTGACAGC	CGAGTCTTGG	AGAGGTACCT	CTTGGAGGCC	AAGGAGGCCG	420
AGAATATCAC	GACGGGCTGT	GCTGAACACT	GCAGCTTGAA	TGAGAATAAT	CACTGTCCCA	480
GACACCAAAG	TTAATTTCTA	TGCCTGGAAG	AGG			513

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GAGGTCGGGC	AGCAGGCCGT	AGAAGTCTGG	CAGGGCCTGG	CCCTGCTGTC	GGAAGCTGTC	60
CTGCGGGGCC	AGGCCCTGTT	GGTCAACTCT	TCCCAGCCGT	GGGAGCCCCT	GCAGCTGCAT	120
GTGGATAAAG	CCGTCACTGG	CCTTCGCAGC	CTCACCCTCT	TGCTTCGGGC	TCTGGGAGCC	180
CAGAAGGAAG	CCATCTCCCC	TCCAGATGCG	GCCTCAGCTG	CTCCACTCCG	AACAATCACT	240
GCTGACACTT	TCCGCAAACCT	CTTCCGAGTC	TACTCCAATT	TCCTCCGGGG	AAAGCTGAAG	300
CTGTACACAG	GGGAGGCCCTG	CAGGACAGGG	GACAGATGAG	GCGGCGGCTC	CCCCCACCAC	360
GCCTCATCTG	TGACAGCCGA	GTCTGGAGA	GGTACCTCTT	GGAGGCCAAG	GAGGCCGAGA	420
ATATCACGAC	GGGCTGTGCT	GAACACTGCA	GCTTGAATGA	GAATAATCAC	TGTCCCAGAC	480
ACCAAAGTTA	ATTCTATGCT	CTGGAAGAGG	ATG			513

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GTCGGGCAGC	AGGCCGTAGA	AGTCTGGCAG	GGCCTGGCCC	TGCTGTCCGA	AGCTGTCTCTG	60
CGGGGCCAGG	CCCTGTTGGT	CAACTCTTCC	CAGCCGTGGG	AGCCCCTGCA	GCTGCATGTG	120
GATAAAGCCG	TCAGTGGCCT	TCGCAGCCTC	ACCACTCTGC	TTCCGGCTCT	GGGAGCCCAG	180
AAGGAAGCCA	TCTCCCCCTC	AGATGCGGCC	TCAGCTGCTC	CACTCCGAAC	AATCACTGCT	240
GACACTTTCC	GCAAACCTTT	CCGAGTCTAC	TCCAATTTCC	TCCGGGGAAA	GCTGAAGCTG	300
TACACAGGGG	AGGCCTGCAG	GACAGGGGAC	AGATGAGGCG	GCGGCTCCCC	CCACCACGCC	360
TCATCTCTGA	CAGCCGAGTC	CTGGAGAGGT	ACCTCTTGGA	GGCCAAGGAG	GCCGAGAATA	420
TCACGACGGG	CTGTGCTGAA	CACTGCAGCT	TGAATGAGAA	TAATCACTGT	CCCAGACACC	480
AAAGTTAATT	TCTATGCCTG	GAAGAGGATG	GAG			513

SUBSTITUTE SHEET (rule 26)

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

CAGGCCCTGT	TGGTCAACTC	TTCCAGCCG	TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	60
GGCGTCAGTG	GCCTTCGCAG	CCTCACCAC	CTGCTTCGGG	CTCTGGGAGC	CCAGAAGGAA	120
GCCATCTCCC	CTCCAGATGC	GGCCTCAGCT	GCTCCACTCC	GAACAATCAC	TGCTGACACT	180
TTCCGCAAAAC	TCTTCCGAGT	CTACTCCAAT	TTCTTCCGGG	GAAAGCTGAA	GCTGTACACA	240
GGGGAGGCCT	GCAGGACAGG	GGACAGATGA	GGCGGCGGCT	CCCCCACCA	CGCCTCATCT	300
GTGACAGCCG	AGTCTGGAG	AGGTACCTCT	TGGAGGCCAA	GGAGGCCGAG	AATATCACGA	360
CGGGCTGTGC	TGAACACTGC	AGCTTGAATG	AGAATAATCA	CTGTCCCAGA	CACCAAAGTT	420
AATTTCTATG	CCTGGAAGAG	GATGGAGGTC	GGGCAGCAGG	CCGTAGAAGT	CTGGCAGGGC	480
CTGGCCCTGC	TGTCGGAAGC	TGTCCTGCGG	GGC			513

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

GCCCTGTGG	TCAACTCTTC	CCAGCCGTGG	GAGCCCCTGC	AGCTGCATGT	GGATAAAGCC	60
GTCAGTGGCC	TTCGCAGCCT	CACCACTCTG	CTTCGGGCTC	TGGGAGCCCA	GAAGGAAGCC	120
ATCTCCCTTC	CAGATGCGGC	CTCAGCTGCT	CCACTCCGAA	CAATCACTGC	TGACACTTTC	180
CGCAAACCTCT	TCCGAGTCTA	CTCCAATTTC	CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	240
GAGGCCTGCA	GGACAGGGGA	CAGATGAGGC	GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	300
ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	360
GCTGTGCTGA	ACACTGCAGC	TTGAATGAGA	ATAATCACTG	TCCCAGACAC	CAAAGTTAAT	420
TTCTATGCCT	GGAAGAGGAT	GGAGCTCGGG	CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	480
GCCCTGCTGT	CGGAAGCTGT	CCTCGGGGCG	CAG			513

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

CTGTTGGTCA	ACTCTTCCCA	GCCGTGGGAG	CCCCTGCAGC	TGCATGTGGA	TAAAGCCGTC	60
AGTGGCCTTC	GCAGCCTCAC	CACTCTGCTT	CGGCTCTTGG	GAGCCCAGAA	GGAAGCCATC	120
TCCCTCCAG	ATGCGGCCTC	AGCTGCTCCA	CTCCGAACAA	TCACTGTGTA	CACCTTCCGC	180
AAACTCTTCC	GAGTCTACTC	CAATTTCCTC	CGGGGAAAGC	TGAAGCTGTA	CACAGGGGAG	240
GCCTGCAGGA	CAGGGGACAG	ATGAGGCGGC	GGCTCCCCC	ACCACGCCTC	ATCTGTGACA	300
GCCGAGTCCT	GGAGAGGTAC	CTCTTGGAGG	CCAAGGAGGC	CGAGAATATC	ACGACGGGCT	360
GTGCTGAACA	CTGCAGCTTG	AATGAGAATA	ATCACTGTCC	CAGACACCAA	AGTTAATTTT	420
TATGCCTGGA	AGAGGATGGA	GGTCGGGCAG	CAGGCCGTAG	AAGTCTGGCA	GGGCTTGGCC	480
CTGCTGTCCG	AAGCTGTCTT	GCGGGGCCAG	GCC			513

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

TTGGTCAACT	CTTCCCAGCC	GTGGGAGCCC	CTGCAGCTGC	ATGTGGATAA	AGCCGTCAGT	60
GGCCTTCGCA	GCCTCACCAC	TCTGCTTCGG	GCTCTGGGAG	CCCAGAAGGA	AGCCATCTCC	120
CCTCCAGATG	CGGCCTCAGC	TGCTCCACTC	CGAACAATCA	CTGCTGACAC	TTTCCGCAAA	180

80

CTCTTCCGAG	TCTACTCCAA	TTTCTCCCG	GGAAAGCTGA	AGCTGTACAC	AGGGGAGGCC	240
TGCAGGACAG	GGGACAGATG	AGGCGGCGGC	TCCCCCACC	ACGCCTCATC	TGTGACAGCC	300
GAGTCCCTGGA	GAGGTACCTC	TTGGAGGCCA	AGGAGGCCGA	GAATATCACG	ACGGGCTGTG	360
CTGAACACTG	CAGCTTGAAT	GAGAATAATC	ACTGTCCCAG	ACACCAAAGT	TAATTTCTAT	420
GCCTGGAAGA	GGATGGAGGT	CGGGCAGCAG	GCCGTAGAAG	TCTGGCAGGG	CCTGGCCCTG	480
CTGTCCGAAG	CTGTCTGCG	GGGCCAGGCC	CTG			513

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

GTCAACTCTT	CCCAGCCGTG	GGAGCCCCTG	CAGCTGCATG	TGGATAAAGC	CGTCAGTGGC	60
CTTCGCAGCC	TCACCACTCT	GCTTCGGGCT	CTGGGAGCCC	AGAAGGAAGC	CATCTCCCCT	120
CCAGATGCGG	CCTCAGCTGC	TCCACTCCGA	ACAATCACTG	CTGACACTTT	CCGCAAACCT	180
TTCCGAGTCT	ACTCCAATTT	CCTCCGGGGA	AAGCTGAAGC	TGTACACAGG	GGAGGCCCTG	240
AGGACAGGGG	ACAGATGAGG	CGGCGGCTCC	CCCCACCACG	CCTCATCTGT	GACAGCCGAG	300
TCCTGGAGAG	GTACCTCTTG	GAGGCCAAGG	AGGCCGAGAA	TATCACCACG	GGCTGTGTCTG	360
AACACTGCAG	CTTGAATGAG	AATAATCACT	GTCCCAGACA	CCAAAGTTAA	TTTCTATGCC	420
TGGAAGAGGA	TGGAGGTCGG	GCAGCAGGCC	GTAGAAGTCT	GGCAGGGCCT	GGCCCTGTCTG	480
TCGGAAGCTG	TCCTGCGGGG	CCAGGCCCTG	TTG			513

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

AACTCTTCCC	AGCCGTGGGA	GCCCCTGCAG	CTGCATGTGG	ATAAAGCCGT	CAGTGGCCTT	60
CGCAGCCTCA	CCACTCTGCT	TCGGGCTCTG	GGAGCCCAGA	AGGAAGCCAT	CTCCCCTCCA	120
GATGCCGCCCT	CAGCTGCTCC	ACTCCGAACA	ATCACTGCTG	ACACTTCCG	CAAACTCTTC	180
CGAGTCTACT	CCAATTTCTT	CCGGGGAAAG	CTGAAGCTGT	ACACAGGGGA	GGCCTGCAGG	240
ACAGGGGACA	GATGAGGCGG	CGGCTCCCCC	CACCACGCCT	CATCTGTGAC	AGCCGAGTCC	300
TGGAGAGGTA	CCTCTTGGAG	GCCAAGGAGG	CCGAGAATAT	CACGACGGGC	TGTGCTGAAC	360
ACTGCAGCTT	GAATGAGAAT	AATCACTGTC	CCAGACACCA	AAGTTAATTT	CTATGCCTGG	420
AAGAGGATGG	AGGTGCGGGC	GCAGGCCGTA	GAAGTCTGGC	AGGGCCTGGC	CCTGCTGTCTG	480
GAAGCTGTCC	TGCGGGGCCA	GGCCTGTGTC	GTC			513

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

TCTTCCAGC	CGTGGGAGCC	CCTGCAGCTG	CATGTGGATA	AAGCCGTCAG	TGGCCTTCGC	60
AGCCTCACCA	CTCTGCTTCG	GGCTCTGGGA	GCCCAGAAGG	AAGCCATCTC	CCCTCCAGAT	120
GCGGCCTCAG	CTGCTCCACT	CCGAACAATC	ACTGCTGACA	CTTTCCGCAA	ACTCTTCCGA	180
GTCTACTCCA	ATTTCTCTCG	GGGAAAGCTG	AAGCTGTACA	CAGGGGAGGC	CTGCAGGACA	240
GGGGACAGAT	GAGGCGGCGG	CTCCCCCACC	CACGCCTCAT	CTGTGACAGC	CGAGTCCTGG	300
AGAGGTACCT	CTTGGAGGCC	AAGGAGGCCG	AGAATATCAC	GACGGGCTGT	GCTGAACACT	360
GCAGCTTGAA	TGAGAATAAT	CACTGTCCCA	GACACCAAAG	TTAATTTCTA	TGCCTGGAAG	420
AGGATGGAGG	TCGGGCAGCA	GGCCGTAGAA	GTCTGGCAGG	GCCTGGCCCT	GCTGTGGGAA	480
GCTGTCTCTG	GGGGCCAGGC	CCTGTTGGTC	AAC			513

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

81

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

TCCCAGCCGT	GGGAGCCCCT	GCAGCTGCAT	GTGGATAAAG	CCGTCAGTGG	CCTTCGCAGC	60
CTCACCACCT	TGCTTCGGGC	TCTGGGAGCC	CAGAAGGAAG	CCATCTCCCC	TCCAGATGCG	120
GCCTCAGCTG	CTCCACTCCG	AACAATCACT	GCTGACACTT	TCCGCAAAC	CTTCCGAGTC	180
TACTCCAATT	TCCTCCGGGG	AAAGCTGAAG	CTGTACACAG	GGGAGGCCCT	GAGGACAGGG	240
GACAGATGAG	GCGGCGGCTC	CCCCCACCAC	GCCTCATCTG	TGACAGCCGA	GTCTTGGAGA	300
GGTACCTCTT	GGAGGCCAAG	GAGGCCGAGA	ATATCACGAC	GGGCTGTGCT	GAACACTGCA	360
GCTTGAATGA	GAATAATCAC	TGTCCCAGAC	ACCAAAGTTA	ATTTCTATGC	CTGGAAGAGG	420
ATGGAGGTCG	GGCAGCAGGC	CGTAGAAGTC	TGGCAGGGCC	TGGCCCTGCT	GTGGAAGACT	480
GTCTTGC	GGG	GCCAGGCCCT	GTTGGTCAAC	TCT		513

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CAGCCGTGGG	AGCCCCCTGCA	GCTGCATGTG	GATAAAGCCG	TCAGTGGCCT	TCGCAGCCTC	60
ACCACTCTGC	TTCGGGCTCT	GGGAGCCCAG	AAGGAAGCCA	TCTCCCCCTC	AGATGCGGCC	120
TCAGCTGTGC	CACTCCGAAC	AATCACTGCT	GACACTTTCC	GCAAACTCTT	CCGAGTCTAC	180
TCCAATTTCC	TCCGGGGAAA	GCTGAAAGTG	TACACAGGGG	AGGCCTGCAG	GACAGGGGAC	240
AGATGAGCGC	GCGGCTCCCC	CCACCACGCC	TCATCTGTGA	CAGCCGAGTC	CTGGAGAGGT	300
ACCTCTTGGG	GGCCAAGGAG	GCCGAGAATA	TCACGACGGG	CTGTGCTGAA	CACTGCAGCT	360
TGAATGAGAA	TAATCACTGT	CCCAGACACC	AAAGTTAATT	TCTATGCCTG	GAAGAGGATG	420
GAGGTGCGGC	AGCAGGCCGT	AGAAGTCTGG	CAGGGCCTGG	CCCTGCTGTC	GGAAGCTGTC	480
CTCGGGGGCC	AGGCCCTGTT	GGTCAACTCT	TCC			513

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

CCGTGGGAGC	CCCTGCAGCT	GCATGTGGAT	AAAGCCGTCA	GTGGCCTTCG	CAGCCTCACC	60
ACTCTGCTTC	GGGCTCTGGG	AGCCCAGAAG	GAAGCCATCT	CCCCCTCCAG	TGCGGCCTCA	120
GCTGCTCCAC	TCCGAACAAT	CACTGCTGAC	ACTTTCGCGA	AACTCTTCCG	AGTCTACTCC	180
AATTTCTCTC	GGGAAAGCT	GAAGCTGTAC	ACAGGGGAGG	CCTGCAGGAC	AGGGGACAGA	240
TGAGGCGGCG	GCTCCCCCCA	CCACGCCCTCA	TCTGTGACAG	CCGAGTCCTG	GAGAGGTACC	300
TCTTGGAGGC	CAAGGAGGCC	GAGAATATCA	CGACGGGCTG	TGCTGAACAC	TGCAGCTTGA	360
ATGAGAATAA	TCACTGTCCC	AGACACCAAA	GTTAATTTCT	ATGCCTGGAA	GAGGATGGAG	420
GTGCGGCAGC	AGGCCGTAGA	AGTCTGGCAG	GGCCTGGGCC	TGCTGTCGGA	AGCTGTCCTG	480
CGGGGCCAGG	CCCTGTTGGT	CAACTCTTCC	CAG			513

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	GCCGTCAGTG	GCCTTCGCAG	CCTCACCAC	60
CTGCTTCGGG	CTCTGGGAGC	CCAGAAGGAA	GCCATCTCCC	CTCCAGATGC	GGCCTCAGCT	120
GCTCCACTCC	GAACAATCAC	TGCTGACACT	TTCCGCAAAC	TCTTCCGAGT	CTACTCCAAT	180
TTCTTCCGGG	GAAAGCTGAA	GCTGTACACA	GGGGAGGCCT	GCAGGACAGG	GGACAGATGA	240
GGCGGCGGCT	CCCCCACCAC	CGCCTCATCT	GTGACAGCCG	AGTCTTGAGG	AGGTACCTCT	300
TGGAGGCCAA	GGAGGCCGAG	AATATCACGA	CGGGCTGTGC	TGAACACTGC	AGCTTGAATG	360
AGAATAATCA	CTGTCCAGA	CACCAAAGTT	AATTTCTATG	CCTGGAAGAG	GATGGAGGTC	420
GGGCAGCAGG	CCGTAGAAGT	CTGGCAGGGC	CTGGCCCTGC	TGTCGGAAGC	TGTCCTGCGG	480
GGCCAGGCCC	TGTTGGTCAA	CTCTTCCCAG	CCG			513

82

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

GAGCCCCTGC	AGCTGCATGT	GGATAAAGCC	GTCAGTGGCC	TTCGCAGCCT	CACCACTCTG	60
CTTCGGGCTC	TGGGAGCCCA	GAAGGAAGCC	ATCTCCCCCTC	CAGATGCGGC	CTCAGCTGCT	120
CCACTCCGAA	CAATCACTGC	TGACACTTTC	CGCAAACTCT	TCCGAGTCTA	CTCCAATTTT	180
CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	GAGGCCTGCA	GGACAGGGGA	CAGATGAGGC	240
GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	300
AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	GCTGTGCTGA	ACACTGCAGC	TTGAATGAGA	360
ATAATCACTG	TCCCAGACAC	CAAAGTTAAT	TTCTATGCCT	GGAAGAGGAT	GGAGGTCGGG	420
CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	GCCCTGCTGT	CGGAAGCTGT	CCTGCGGGGC	480
CAGGCCCTGT	TGGTCAACTC	TTCCAGCCG	TGG			513

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

CTTCGGGCTC	TGGGAGCCCA	GAAGGAAGCC	ATCTCCCCCTC	CAGATGCGGC	CTCAGCTGCT	60
CCACTCCGAA	CAATCACTGC	TGACACTTTC	CGCAAACTCT	TCCGAGTCTA	CTCCAATTTT	120
CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	GAGGCCTGCA	GGACAGGGGA	CAGATGAGGC	180
GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	240
AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	GCTGTGCTGA	ACACTGCAGC	TTGAATGAGA	300
ATAATCACTG	TCCCAGACAC	CAAAGTTAAT	TTCTATGCCT	GGAAGAGGAT	GGAGGTCGGG	360
CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	GCCCTGCTGT	CGGAAGCTGT	CCTGCGGGGC	420
CAGGCCCTGT	TGGTCAACTC	TTCCAGCCG	TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	480
GCCGTCACTG	GCCTTCGCAG	CCTCACCCT	CTG			513

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CGGGCTCTGG	GAGCCAGAA	GGAAGCCATC	TCCCCTCCAG	ATGCGGCCTC	AGCTGTCCA	60
CTCCGAACAA	TCACTGCTGA	CACCTTCCGC	AAACTCTTCC	GAGTCTACTC	CAATTTCTCT	120
CGGGGAAAGC	TGAAGCTGTA	CACAGGGGAG	GCCTGCAGGA	CAGGGGACAG	ATGAGGCGGC	180
GGCTCCCCC	ACCACGCCCT	ATCTGTGACA	GCCGAGTCCT	GGAGAGGTAC	CTCTTGGAGG	240
CCAAGGAGGC	CGAGAATATC	ACGACGGGCT	GTGCTGAACA	CTGCAGCTTG	AATGAGAATA	300
ATCACTGTCC	CAGACACCAA	AGTTAATTTT	TATGCCTGGA	AGAGGATGGA	GGTCGGGCAG	360
CAGGCCGTAG	AAGTCTGGCA	GGGCCTGGCC	CTGCTGTGCG	AAGCTGTCCT	GCGGGGCCAG	420
GCCCTGTTGG	TCAACTCTTC	CCAGCCGTGG	GAGCCCCTGC	AGCTGCATGT	GGATAAAGCC	480
GTCAGTGGCC	TTCGCAGCCT	CACCACTCTG	CTT			513

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GCTCTGGGAG	CCCAGAAGGA	AGCCATCTCC	CCTCCAGATG	CGGCCTCAGC	TGCTCCACTC	60
CGAAACAATCA	CTGCTGACAC	TTTCCGCAA	CTCTTCCGAG	TCTACTCCAA	TTTCTCTCCG	120
GGAAGCTGA	AGCTGTACAC	AGGGGAGGCC	TGCAGGACAG	GGGACAGATG	AGGCGGCGGC	180

83

TCCCCCACC	ACGCCTCATC	TGTGACAGCC	GAGTCCTGGA	GAGGTACCTC	TTGGAGGCCA	240
AGGAGGCCGA	GAATATCACG	ACGGGCTGTG	CTGAACACTG	CAGCTTGAAT	GAGAATAATC	300
ACTGTCCCAG	ACACCAAAGT	TAATTTCTAT	GCCTGGAAGA	GGATGGAGGT	CGGGCAGCAG	360
GCCGTAGAAG	TCTGGCAGGG	CCTGGCCCTG	CTGTGGAAG	CTGTCTGCG	GGGCCAGGCC	420
CTGTTGGTCA	ACTCTTCCCA	GCCGTGGGAG	CCCCTGCAGC	TGCATGTGGA	TAAAGCCGTC	480
AGTGGCCTTC	GCAGCTCAC	CACTCTGCTT	CGG			513

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

CTGGGAGCCC	AGAAGGAAGC	CATCTCCCCT	CCAGATGCGG	CCTCAGCTGC	TCCACTCCGA	60
ACAATCACTG	CTGACACTTT	CCGCAAACCTC	TTCCGAGTCT	ACTCCAATTT	CCTCCGGGGA	120
AAGCTGAAGC	TGTACACAGG	GGAGGCCTGC	AGGACAGGGG	ACAGATGAGG	CGGCGGCTCC	180
CCCCACCACG	CCTCATCTGT	GACAGCCGAG	TCCTGGAGAG	GTACCTCTTG	GAGGCCAAGG	240
AGGCCGAGAA	TATCACGACG	GGCTGTGCTG	AACACTGCAG	CTTGAATGAG	AATAATCACT	300
GTCCCAAGACA	CCAAAGTTAA	TTTCTATGCC	TGGAAGAGGA	TGGAGGTCGG	GCAGCAGGCC	360
GTAGAAGTCT	GGCAGGGCCT	GGCCCTGCTG	TCGGAAGCTG	TCCTGCGGGG	CCAGGCCCTG	420
TTGGTCAACT	CTTCCCAGCC	GTGGGAGCCC	CTGCAGCTGC	ATGTGGATAA	AGCCGTCAGT	480
GGCCTTCGCA	GCCTACCCAC	TCTGCTTCGG	GTC			513

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GGAGCCCAGA	AGGAAGCCAT	CTCCCCTCCA	GATGCGGCCT	CAGCTGCTCC	ACTCCGAACA	60
ATCACTGCTG	ACACTTTCCG	CAAACCTCTT	CGAGTCTACT	CCAATTTCTT	CCGGGGAAAG	120
CTGAAGCTGT	ACACAGGGGA	GGCCTGCAGG	ACAGGGGACA	GATGAGGCGG	CGGCTCCCCC	180
CACCACGCCT	CATCTGTGAC	AGCCGAGTCC	TGGAGAGGTA	CCTCTTGAG	GCCAAGGAGG	240
CCGAGAAATAT	CACGACGGGC	TGTGCTGAAC	ACTGCAGCTT	GAATGAGAAT	AATCACTGTC	300
CCAGACACCA	AAGTTAATTT	CTATGCCTGG	AAGAGGATGG	AGGTCGGGCA	GCAGGCCGTA	360
GAAGTCTGGC	AGGGCCTGGC	CCTGTCTGTC	GAAGCTGTCC	TGCGGGGCCA	GGCCCTGTTG	420
GTCAACTCTT	CCCAGCCGTG	GGAGCCCTG	CAGCTGCATG	TGGATAAAGC	CGTCAGTGCC	480
CTTCGCAGCC	TCACCACTCT	GCTTCGGGCT	CTG			513

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

GCCCAGAAGG	AAGCCATCTC	CCCTCCAGAT	GCGGCCTCAG	CTGCTCCACT	CCGAACAATC	60
ACTGCTGACA	CTTTCCGCAA	ACTCTTCCGA	GTCTACTCCA	ATTTCCTCCG	GGGAAAGCTG	120
AAGCTGTACA	CAGGGGAGGC	CTGCAGGACA	GGGGACAGAT	GAGGCGGCGG	CTCCCCCAC	180
CACGCCTCAT	CTGTGACAGC	CGAGTCCTGG	AGAGGTACCT	CTTGGAGGCC	AAGGAGGCCG	240
AGAATATCAC	GACGGGCTGT	GCTGAACACT	GCAGCTTGAA	TGAGAATAAT	CACTGTCCCA	300
GACACCAAAG	TAAATTTCTA	TGCCTGGAAG	AGGATGGAGG	TCGGGCAGCA	GGCCGTAGAA	360
GTCTGGCAGG	GCCTGGCCCT	GCTGTGCGAA	GCTGTCTGCG	GGGGCCAGGC	CCTGTTGGTC	420
AACTCTTCCC	AGCCGTGGGA	GCCCTGCAG	CTGCATGTGG	ATAAAGCCGT	CAGTGGCCTT	480
CGCAGCCTCA	CCACTCTGCT	TCGGGCTCTG	GGA			513

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

84

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CAGAAGGAAG	CCATCTCCCC	TCCAGATGCG	GCCTCAGCTG	CTCCACTCCG	AACAATCACT	60
GCTGACACTT	TCCGCAAACT	CTTCCGAGTC	TACTCCAATT	TCCTCCGGGG	AAAGCTGAAG	120
CTGTACACAG	GGGAGGCCCTG	CAGGACAGGG	GACAGATGAG	GCGGCGGCTC	CCCCCACCAC	180
GCCTCATCTG	TGACAGCCGA	GTCCTGGAGA	GGTACCTCTT	GGAGGCCAAG	GAGGCCGAGA	240
ATATCACGAC	GGGCTGTGCT	GAACACTGCA	GCTTGAATGA	GAATAATCAC	TGTCCCAGAC	300
ACCAAAGTTA	ATTTCTATGC	CTGGAAGAGG	ATGGAGGTCG	GGCAGCAGGC	CGTAGAAGTC	360
TGGCAGGGCC	TGGCCCTGCT	GTCGGAAGCT	GTCTGCGGG	GCCAGGCCCT	GTTGGTCAAC	420
TCTTCCCAGC	CGTGGGAGCC	CCTGCAGCTG	CATGTGGATA	AAGCCGTCAG	TGGCCTTCGC	480
AGCCTCACCA	CTCTGCTTCG	GGCTCTGGGA	GCC			513

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

AAGGAAGCCA	TCTCCCCTCC	AGATGCGGCC	TCAGCTGCTC	CACTCCGAAC	AATCACTGCT	60
GACACTTTCC	GCAAACTCTT	CCGAGTCTAC	TCCAATTTC	TCCGGGGAAA	GCTGAAGCTG	120
TACACAGGGG	AGGCCTGCAG	GACAGGGGAC	AGATGAGGCG	GCGGCTCCCC	CCACCACGCC	180
TCATCTGTGA	CAGCCGAGTC	CTGGAGAGGT	ACCTCTTGGA	GGCCAAAGGAG	GCCGAGAATA	240
TCACGACGGG	CTGTGCTGAA	CACTGCAGCT	TGAATGAGAA	TAATCACTGT	CCCAGACACC	300
AAAGTTAATT	TCTATGCCCTG	GAAGAGGATG	GAGGTCGGGC	AGCAGGCCGT	AGAAGTCTGG	360
CAGGGCCTGG	CCCTGCTGTC	GGAAGCTGTC	CTGCGGGGCC	AGGCCCTGTT	GGTCAACTCT	420
TCCCAGCCGT	GGGAGCCCCCT	GCAGCTGCAT	GTGGATAAAG	CCGTCAGTGG	CCTTCGCAGC	480
CTCACCACTC	TGCTTCGGGC	TCTGGGAGCC	CAG			513

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

GAAGCCATCT	CCCCTCCAGA	TGCGGCCTCA	GCTGCTCCAC	TCCGAACAAT	CACTGCTGAC	60
ACTTTCCGCA	AACTCTTCCG	AGTCTACTCC	AATTTCCCTC	GGGGAAAGCT	GAAGCTGTAC	120
ACAGGGGAGG	CCTGCAGGAC	AGGGGACAGA	TGAGGCGGCG	GCTCCCCCCA	CCACGCCCTCA	180
TCTGTGACAG	CCGAGTCTCTG	GAGAGGTACC	TCTTGGAGGC	CAAGGAGGCC	GAGAATATCA	240
CGACGGGCTG	TGCTGAACAC	TGCAGCTTGA	ATGAGAATAA	TCACTGTCCC	AGACACCAAA	300
GTTAATTTCT	ATGCCTGGAA	GAGGATGGAG	GTGCGGCAGC	AGGCCCTAGA	AGTCTGGCAG	360
GGCCTGGCCC	TGCTGTCCGA	AGCTGTCTCTG	CGGGGCCAGG	CCCTGTTGGT	CAACTCTTCC	420
CAGCCGTGGG	AGCCCCTGCA	GCTGCATGTG	GATAAAGCCG	TCAGTGGCCT	TCGCAGCCTC	480
ACCACCTCTG	TTGCGGCTCT	GGGAGCCCAG	AAG			513

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GCCATCTCCC	CTCCAGATGC	GGCCTCAGCT	GCTCCACTCC	GAACAATCAC	TGCTGACACT	60
TTCCGCAAAAC	TCTTCCGAGT	CTACTCCAAT	TTCCCTCCGG	GAAAGCTGAA	GCTGTACACA	120
GGGGAGGCCCT	GCAGGACAGG	GGACAGATGA	GGCGGCGGCT	CCCCCACCAC	CGCCTCATCT	180
GTGACAGCCG	AGTCCTGGAG	AGGTACCTCT	TGGAGGCCAA	GGAGGCCGAG	AATATCACGA	240
CGGGCTGTGC	TGAACACTGC	AGCTTGAATG	AGAATAATCA	CTGTCCCAGA	CACCAAAGTT	300
AATTTCTATG	CCTGGAAGAG	GATGGAGGTC	GGCAGCAGG	CCGTAGAAGT	CTGGCAGGGC	360
CTGGCCCTGC	TGTCGGAAGC	TGTCCTGCCG	GGCCAGGCC	TSTTGGTCAA	CTCTTCCCAG	420
CCGTGGGAGC	CCCTGCAGCT	GCATGTGGAT	AAAGCCGTCA	GTGGCCCTTC	CAGCCTCACC	480
ACTCTGCTTC	GGGCTCTGGG	AGCCCAGAAG	GAA			513

85

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

ATCTCCCTC	CAGATGCGGC	CTCAGCTGCT	CCACTCCGAA	CAATCACTGC	TGACACTTTC	60
CGCAAACCT	TCCGAGTCTA	CTCCAATTTC	CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	120
GAGGCCGTG	GGACAGGGGA	CAGATGAGGC	GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	180
ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	240
GCTGTGCTGA	ACACTGCAGC	TTGAATGAGA	ATAATCACTG	TCCCAGACAC	CAAAGTTAAT	300
TTCTATGCCT	GGAAGAGGAT	GGAGGTCGGG	CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	360
GCCCTGCTGT	CGGAAGCTGT	CCTGCGGGGC	CAGGCCCTGT	TGGTCAACTC	TTCCCAGCCG	420
TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	GCCGTCAGTG	GCCTTCGCAG	CCTCACCAC	480
CTGCTTCGGG	CTCTGGGAGC	CCAGAAGGAA	GCC			513

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

TCCCCTCCAG	ATGCGGCCTC	AGCTGCTCCA	CTCCGAACAA	TCACTGCTGA	CACTTTCCGC	60
AAACTCTTCC	GAGTCTACTC	CAATTTCCTC	CGGGGAAAGC	TGAAGCTGTA	CACAGGGGAG	120
GCCTGCAGGA	CAGGGGACAG	ATGAGGCGGC	GGCTCCCCCC	ACCACGCCTC	ATCTGTGACA	180
GCCGAGTCCT	GGAGAGGTAC	CTCTTGGAGG	CCAAGGAGGC	CGAGAATATC	ACGACGGGCT	240
GTGCTGAACA	CTGCAGCTTG	AATGAGAATA	ATCACTGTCC	CAGACACCAA	AGTTAATTTC	300
TATGCCTGGA	AGAGGATGGA	GGTCGGGCAG	CAGGCCGTAG	AAGTCTGGCA	GGGCCCTGGC	360
CTGCTGTCCG	AAGCTGTCCT	GCGGGGCCAG	GCCCTGTTGG	TCAACTCTTC	CCAGCCGTGG	420
GAGCCCTGCG	AGCTGCATGT	GGATAAAGCC	GTCAGTGGCC	TTCGCAGCCT	CACCACTCTG	480
CTTCGGGCTC	TGGGAGCCCC	GAAGGAAGCC	ATC			513

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

CCTCCAGATG	CGGCCTCAGC	TGCTCCACTC	CGAACAATCA	CTGCTGACAC	TTTCCGCAAA	60
CTCTTCCGAG	TCTACTCCAA	TTTCTCTCCG	GGAAAGCTGA	AGCTGTACAC	AGGGGAGGCC	120
TGCAGGACAG	GGGACAGATG	AGGCGGCGGC	TCCCCCACC	ACGCCTCATC	TGTGACAGCC	180
GAGTCCTGGA	GAGGTACCTC	TTGGAGGCCA	AGGAGGCCGA	GAATATCACG	ACGGGCTGTG	240
CTGAACACTG	CAGCTTGAAT	GAGAATAATC	ACTGTCCCAG	ACACCAAAGT	TAATTTCTAT	300
GCCTGGAAGA	GGATGGAGGT	CGGGCAGCAG	GCCGTAGAAG	TCTGGCAGGG	CCTGGCCCTG	360
CTGTGCGAAG	CTGTCTGCG	GGGCCAGGCC	CTGTTGGTCA	ACTCTTCCCA	GCCGTGGGAG	420
CCCCTGCAGC	TGCATGTGGA	TAAAGCCGTC	AGTGGCCTTC	GCAGCCTCAC	CACCTGTCTT	480
CGGGCTCTGG	GAGCCAGAA	GGAAGCCATC	TCC			513

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

CCAGATGCGG	CCTCAGCTGC	TCCACTCCGA	ACAATCACTG	CTGACACTTT	CCGCAAACTC	60
TTCCGAGTCT	ACTCCAATTT	CCTCCGGGGA	AAGCTGAAGC	TGTACACAGG	GGAGGCCTGC	120
AGGACAGGGG	ACAGATGAGG	CGGCGGCTCC	CCCCACCACG	CCTCATCTGT	GACAGCCGAG	180

86

TCCTGGAGAG	GTACCTCTTG	GAGGCCAAGG	AGGCCGAGAA	TATCACGACG	GGCTGTGCTG	240
AACACTGCAG	CTTGAATGAG	AATAATCACT	GTCCCCAGACA	CCAAAGTTAA	TTTCTATGCC	300
TGGAAGAGGA	TGGAGGTCGG	GCAGCAGGCC	GTAGAAGTCT	GGCAGGGCCT	GGCCCTGCTG	360
TCGGAAGCTG	TCCTGCGGGG	CCAGGCCCTG	TTGGTCAACT	CTTCCCAGCC	GTGGGAGCCC	420
CTGCAGCTGC	ATGTGGATAA	AGCCGTCAGT	GGCCTTCGCA	GCCTCACCAC	TCTGCTTCGG	480
GCTCTGGGAG	CCCAGAAGGA	AGCCATCTCC	CCT			513

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

GATGCGGCCT	CAGCTGCTCC	ACTCCGAACA	ATCACTGCTG	ACACTTTCGG	CAAACCTCTC	60
CGAGTCTACT	CCAATTTCCT	CCGGGGAAG	CTGAAGCTGT	ACACAGGGGA	GGCCTGCAGG	120
ACAGGGGACA	GATGAGGCGG	CGGCTCCCCC	CACCAAGCCT	CATCTGTGAC	AGCCGAGTCC	180
TGGAGAGGTA	CCTCTGGAG	GCCAAGGAGG	CCGAGAATAT	CACGACGGGC	TGTGCTGAAC	240
ACTGCAGCTT	GAATGAGAAT	AATCACTGTC	CCAGACACCA	AAGTTAATTT	CTATGCCTGG	300
AAGAGGATGG	AGGTCGGGCA	GCAGGCCGTA	GAAGTCTGGC	AGGGCCTGGC	CCTGCTGTCTG	360
GAAGCTGTCC	TGCGGGGCCA	GGCCTGTGTG	GTCAACTCTT	CCCAGCCGTG	GGAGCCCTTG	420
CAGCTGCATG	TGGATAAAGC	CGTCAGTGGC	CTTCGCAGCC	TCACCACTCT	GCTTCGGGCT	480
CTGGGAGCCC	AGAAGGAAGC	CATCTCCCTC	CCA			513

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GCGGCCTCAG	CTGCTCCACT	CCGAACAATC	ACTGCTGACA	CTTCCGCAA	ACTCTCCGA	60
GTCTACTCCA	ATTTCTCTCC	GGGAAAGCTG	AAGCTGTACA	CAGGGGAGGC	CTGCAGGACA	120
GGGGACAGAT	GAGGCGGGCG	CTCCCCCACC	CACGCCCTCAT	CTGTGACAGC	CGAGTCCTGG	180
AGAGGTACCT	CTTGGAGGCC	AAGGAGGCCG	AGAATATCAC	GACGGGCTGT	GCTGAACACT	240
GCAGCTTGAA	TGAGAATAAT	CACTGTCCCA	GACACCAAAG	TTAATTTCTA	TGCCTGGAAG	300
AGGATGGAGG	TCGGGCAGCA	GGCCGTAGAA	GTCTGGCAGG	GCCTGGCCCT	GCTGTCCGAA	360
GCTGTCTCTG	GGGGCCAGGC	CCTGTTGGTC	AACTCTTCCC	AGCCGTGGGA	GCCCCCTGCAG	420
CTGCATGTGG	ATAAGCCGT	CAGTGGCCTT	CGCAGCCTCA	CCACTCTGCT	TCGGGCTCTG	480
GGAGCCCAGA	AGGAAGCCAT	CTCCCCCTCA	GAT			513

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

GCCTCAGCTG	CTCCACTCCG	AACAATCACT	GCTGACACTT	TCCGCAAAC	CTTCCGAGTC	60
TACTCCAATT	TCCTCGGGGG	AAAGCTGAAG	CTGTACACAG	GGGAGGCCTG	CAGGACAGGG	120
GACAGATGAG	GCGGCGGCTC	CCCCCACCAC	GCCTCATCTG	TGACAGCCGA	GTCTTGGAGA	180
GGTACCTCTT	GGAGGCCAAG	GAGGCCGAGA	ATATCACGAC	GGGCTGTGCT	GAACACTGCA	240
GCTTGAATGA	GAATAATCAC	TGTCCCAGAC	ACCAAAGTTA	ATTTCTATGC	CTGGAAGAGG	300
ATGGAGGTCG	GGCAGCAGGC	CGTAGAAGTC	TGGCAGGGCC	TGGCCCTGCT	GTCCGGAAGCT	360
GTCTTGCGGG	GCCAGGCCCT	GTTGGTCAAC	TCTTCCCAGC	CGTGGGAGCC	CCTGCAGCTG	420
CATGTGGATA	AAGCCGTACG	TGGCCTTCGC	AGCCTCACCA	CTCTGCTTCG	GGCTCTGGGA	480
GCCCAGAAGG	AAGCCATCTC	CCCTCCAGAT	GCG			513

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

87

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

TCAGCTGCTC	CACTCCGAAC	AATCACTGCT	GACACTTTCC	GCAAACTCTT	CCGAGTCTAC	60
TCCAATTTC	TCCGGGGAAA	GCTGAAGCTG	TACACAGGGG	AGGCCTGCAG	GACAGGGGAC	120
AGATGAGGCG	GCGGCTCCCC	CCACCACGCC	TCATCTGTGA	CAGCCGAGTC	CTGGAGAGGT	180
ACCTCTTGGA	GGCCAAGGAG	GCCGAGAATA	TCACGACGGG	CTGTGCTGAA	CACTGCAGCT	240
TGAATGAGAA	TAATCACTGT	CCCAGACACC	AAAGTTAATT	TCTATGCCTG	GAAGAGGATG	300
GAGGTCGGGC	AGCAGGCCGT	AGAAGTCTGG	CAGGGCCTGG	CCCTGCTGTC	GGAAGCTGTC	360
CTGCGGGGCC	AGGCCCTGTT	GGTCAACTCT	TCCCAGCCGT	GGGAGCCCTT	GCAGCTGCAT	420
GTGGATAAAG	CCGTCAGTGG	CCTTCGCAGC	CTCACCACCT	TGCTTCGGGC	TCTGGGAGCC	480
CAGAAGGAAG	CCATCTCCCC	TCCAGATGCG	GCC			513

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

GCTGCTCCAC	TCCGAACAAT	CACTGCTGAC	ACTTTCCGCA	AACTCTTCCG	AGTCTACTCC	60
AATTTCTCTC	GGGGAAAGCT	GAAGCTGTAC	ACAGGGGAGG	CCTGCAGGAC	AGGGGACAGA	120
TGAGGCGGCG	GCTCCCCCA	CCACGCCTCA	TCTGTGACAG	CCGAGTCCTG	GAGAGGTACC	180
TCTTGGAGGC	CAAGGAGGCC	GAGAATATCA	CGACGGGCTG	TGCTGAACAC	TGCAGCTTGA	240
ATGAGAATAA	TCACTGTCCC	AGACACCAAA	GTTAATTTCT	ATGCCTGGAA	GAGGATGGAG	300
GTCGGGCAGC	AGGCCGTAGA	AGTCTGGCAG	GGCCTGGCCC	TGCTGTCCGA	AGCTGTCTCT	360
CGGGGCCAGG	CCCTGTTGGT	CAACTCTTCC	CAGCCGTGGG	AGCCCCTGCA	GCTGCATGTG	420
GATAAAGCCG	TCAGTGGCCT	TCGCAGCCTC	ACCACTCTGC	TTCGGGCTCT	GGGAGCCCA	480
AAGGAAGCCA	TCTCCCTTCC	AGATGCGGCC	TCA			513

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

GCTCCACTCC	GAACAATCAC	TGCTGACACT	TTCCGCAAAC	TCTTCCGAGT	CTACTCCAAT	60
TTCTCTCCGG	GAAAGCTGAA	GCTGTACACA	GGGGAGGCCT	GCAGGACAGG	GGACAGATGA	120
GGCGGGCGGT	CCCCCACCAC	CGCCTCATCT	GTGACAGCCG	AGTCCTGGAG	AGGTACTCTT	180
TGGAGGCCAA	GGAGGCCGAG	AATATCACGA	CGGGCTGTGC	TGAACACTGC	AGCTTGAATG	240
AGAATAATCA	CTGTCCCAGA	CACCAAAGTT	AATTCTATG	CCTGGAAGAG	GATGGAGGTC	300
GGGCAGCAGG	CCGTAGAAGT	CTGGCAGGGC	CTGGCCCTGC	TGTCGGAAGC	TGTCCTGCGG	360
GGCCAGGCC	TGTTGGTCAA	CTCTTCCCAG	CCGTGGGAGC	CCCTGCAGCT	GCATGTGGAT	420
AAAGCCGTCA	TGGGCCTTCG	CAGCCTCACC	ACTCTGCTTC	GGGCTCTGGG	AGCCCAGAAG	480
GAAGCCATCT	CCCCCTCCAG	TGCGGCCTCA	GCT			513

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

CCACTCCGAA	CAATCACTGC	TGACACTTTC	CGCAAACCTCT	TCCGAGTCTA	CTCCAATTTT	60
CTCCGGGGAA	AGCTGAAGCT	GTACACAGGG	GAGGCCTGCA	GGACAGGGGA	CAGATGAGGC	120
GGCGGCTCCC	CCCACCACGC	CTCATCTGTG	ACAGCCGAGT	CCTGGAGAGG	TACCTCTTGG	180
AGGCCAAGGA	GGCCGAGAAT	ATCACGACGG	GCTGTGCTGA	ACACTGCAGC	TTGAATGAGA	240
ATAATCACTG	TCCCAGACAC	CAAAGTTAAT	TTCTATGCCT	GGAAGAGGAT	GGAGGTCGGG	300
CAGCAGGCCG	TAGAAGTCTG	GCAGGGCCTG	GCCCTGTGCT	CGGAAGCTGT	CCTGCGGGGC	360
CAGGCCCTGT	TGCTCAACTC	TTCCAGCCG	TGGGAGCCCC	TGCAGCTGCA	TGTGGATAAA	420
GCCGTCAGTG	GCCTTCGCAG	CCTCACCAC	CTGCTTCGGG	CTCTGGGAGC	CCAGAAGGAA	480
GCCATCTCCC	CTCCAGATGC	GGCCTCAGCT	GCT			513

88

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

CTCCGAACAA	TCACTGCTGA	CACTTTCCGC	AAACTCTTCC	GAGTCTACTC	CAATTTCTCTC	60
CGGGGAAAGC	TGAAGCTGTA	CACAGGGGAG	GCCTGCAGGA	CAGGGGACAG	ATGAGGCGGC	120
GGCTCCCCC	ACCACGCCTC	ATCTGTGACA	GCCGAGTCCT	GGAGAGGTAC	CTCTTGGAGG	180
CCAAGGAGGC	CGAGAATATC	ACGACGGGCT	GTGCTGAACA	CTGCAGCTTG	AATGAGAATA	240
ATCACTGTCC	CAGACACCAA	AGTTAATTTT	TATGCCTGGA	AGAGGATGGA	GGTCGGGCAG	300
CAGGCCGTAG	AAGTCTGGCA	GGGCCTGGCC	CTGCTGTCCG	AAGCTGTCTT	GCGGGGCCAG	360
GCCCTGTTGG	TCAACTCTTC	CCAGCCGTGG	GAGCCCTTGC	AGCTGCATGT	GGATAAAGCC	420
GTCAGTGGCC	TTCGCAGCCT	CACCACTCTG	CTTCGGGCTC	TGGGAGCCCA	GAAGGAAGCC	480
ATCTCCCCTC	CAGATGCGGC	CTCAGCTGCT	CCA			513

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

CGAACAATCA	CTGCTGACAC	TTTCCGCAAA	CTCTTCCGAG	TCTACTCCAA	TTTCCTCCGG	60
GGAAAGCTGA	AGCTGTACAC	AGGGGAGGCC	TGCAGGACAG	GGGACAGATG	AGGCGGCGGC	120
TCCCCCACC	ACGCCTCATC	TGTGACAGCC	GAGTCTTGGA	GAGGTACCTC	TTGGAGGCCA	180
AGGAGGCCGA	GAATATCACG	ACGGGCTGTG	CTGAACACTG	CAGCTTGAAT	GAGAATAATC	240
ACTGTCCAG	ACACCAAAGT	TAATTTCTAT	GCCTGGAAGA	GGATGGAGGT	CGGGCAGCAG	300
GCCGTAGAAG	TCTGGCAGGG	CCTGGCCCTG	CTGTGGAAG	CTGTCTCGCG	GGGCCAGGCC	360
CTGTTGGTCA	ACTCTTCCCA	GCCGTGGGAG	CCCCTGCAGC	TGCATGTGGA	TAAAGCCGTC	420
AGTGGCCTTC	GCAGCCTCAC	CACTCTGCTT	CGGGCTCTGG	GAGCCCAGAA	GGAAGCCATC	480
TCCCTCCAG	ATGCGGCCTC	AGCTGTCTCA	CTC			513

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

ACAATCACTG	CTGACACTTT	CCGCAAACCT	TTCCGAGTCT	ACTCCAATTT	CCTCCGGGGA	60
AAGCTGAAGC	TGTACACAGG	GGAGGCCTGC	AGGACAGGGG	ACAGATGAGG	CGGCGGCTCC	120
CCCCACCACG	CCTCATCTGT	GACAGCCGAG	TCCTGGAGAG	GTACCTCTTG	GAGGCCAAGG	180
AGGCCGAGAA	TATCACGACG	GGCTGTGCTG	AACACTGCAG	CTTGAATGAG	AATAATCACT	240
GTCCAGACAA	CCAAAGTTAA	TTTCTATGCC	TGGAAGAGGA	TGGAGGTCGG	GCAGCAGGCC	300
GTAGAAGTCT	GGCAGGGCCT	GGCCCTGCTG	TCGGAAGCTG	TCCTGCGGGG	CCAGGCCCTG	360
TTGGTCAACT	CTTCCCAGCC	GTGGGAGCCC	CTGCAGCTGC	ATGTGGATAA	AGCCGTCAGT	420
GGCCTTCGCA	GCCTCACCAC	TCTGCTTCGG	GCTCTGGGAG	CCCAGAAGGA	AGCCATCTCC	480
CCTCCAGATG	CGGCCTCAGC	TGCTCCACTC	CGA			513

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 501 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

GCCCCACCAC	GCCTCATCTG	TGACAGCCGA	GTCCTGGAGA	GGTACCTCTT	GGAGGCCAAG	60
GAGGCCGAGA	ATATCACGAC	GGGCTGTGCT	GAACACTGCA	GCTTGAATGA	GAATATCACT	120
GTCCAGACAA	CCAAAGTTAA	TTTCTATGCC	TGGAAGAGGA	TGGAGGTCGG	GCAGCAGGCC	180

89

```

GTAGAAGTCT GGCAGGGCCT GGCCCTGCTG TCGGAAGCTG TCCTGCGGGG CCAGGCCCTG 240
TTGGTCAACT CTTCCAGCC GTGGGAGCCC CTGCAGCTGC ATGTGGATAA AGCCGTCAGT 300
GGCCTTCGCA GCCTCACCAC TCTGCTTCGG GCTCTGGGAG CCCAGAAGGA AGCCATCTCC 360
CCTCCAGATG CGGCCTCAGC TGCTCCACTC CGAACAATCA CTGCTGACAC TTTCCGCAAA 420
CTCTTCCGAG TCTACTCAA TTTCTCCGG GGAAGCTGA AGCTGTACAC AGGGGAGGCC 480
TGCAGGACAG GGGACAGATG A 501

```

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

```

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
 1          5          10          15
Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
 20          25          30
Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
 35          40          45
Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
 50          55          60
Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
 65          70          75          80
Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
 85          90          95
Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
100          105          110
Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
115          120          125
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
130          135          140
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145          150          155          160
Cys Arg Thr Gly Asp Arg
165

```

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

```

Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu
 1          5          10          15
Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser
 20          25          30
Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro
 35          40          45
Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg
 50          55          60
Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile
 65          70          75          80
Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala
 85          90          95
Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly
100          105          110
Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg Gly
115          120          125
Gly Gly Ser Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu
130          135          140
Arg Tyr Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys
145          150          155          160
Ala Glu His Cys Ser Leu Asn Glu Asn Ile
165          170

```

(2) INFORMATION FOR SEQ ID NO:123:

90

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: None
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Gly Gly Gly Ser
1

- (2) INFORMATION FOR SEQ ID NO:124:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: None
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Gly Gly Gly Ser Gly Gly Gly Ser
1 5

- (2) INFORMATION FOR SEQ ID NO:125:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: None
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
1 5 10

- (2) INFORMATION FOR SEQ ID NO:126:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: None
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Ser Gly Gly Ser Gly Gly Ser
1 5

- (2) INFORMATION FOR SEQ ID NO:127:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: None
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Glu Phe Gly Asn Met
1 5

- (2) INFORMATION FOR SEQ ID NO:128:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids

91

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Glu Phe Gly Gly Asn Met
1 5

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Glu Phe Gly Gly Asn Gly Gly Asn Met
1 5

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Gly Gly Ser Asp Met Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

GGCGCCCCAT GGACAATCAC TGCTGAC

27

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

TCTGTCCCCT GTCCT

15

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

SUBSTITUTE SHEET (rule 26)

92

GCGCGCAAGC TTATTATCGG AGTGGAGCAG CTGAGGCCGC ATC

43

(2) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

GCCCCACCAC GCCTCATCTG T

21

WHAT IS CLAIMED IS: 93

1. A human EPO receptor agonist polypeptide,
comprising a modified EPO amino acid sequence of the
5 Formula:

	AlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys	
	10	20
10	GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr	
	30	40
	ValProAspThrLysValAsnPheTyrAlaTrpLysArgMetGluValGlyGlnGlnAla	
15	50	60
	ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu	
	70	80
20	LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer	
	90	100
	GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer	
	110	120
25	ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys	
	130	140
	LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla	
30	150	160
	CysArgThrGlyAspArg	
	SEQ ID NO:121	
	166	

wherein optionally 1-6 amino acids from the N-
35 terminus and 1-5 from the C-terminus can be deleted
from said EPO receptor agonist polypeptide;

wherein the N-terminus is joined to the C-terminus
directly or through a linker capable of joining the
40 N-terminus to the C-terminus and having new C- and N-
termini at amino acids;

23-24	48-49	111-112
24-25	50-51	112-113
25-26	51-52	113-114
26-27	52-53	114-115
27-28	53-54	115-116
28-29	54-55	116-117
29-30	55-56	117-118
30-31	56-57	118-119

	⁹⁴	
31-32	57-58	119-120
32-33	77-78	120-121
33-34	78-79	121-122
34-35	79-80	122-123
35-36	80-81	123-124
36-37	81-82	124-125
37-38	82-83	125-126
38-39	84-85	126-127
40-41	85-86	127-128
41-42	86-87	128-129
43-44	87-88	129-130
44-45	88-89	130-131
45-46	108-109	131-132
46-47	109-110	respectively; and
47-48	110-111	

said EPO receptor agonist polypeptide may optionally be immediately preceded by (methionine⁻¹), (alanine⁻¹) or (methionine⁻², alanine⁻¹).

5

2. The EPO receptor agonist polypeptide, as recited in claim 1, wherein said linker is selected from the group consisting of;

- 10 GlyGlyGlySer SEQ ID NO:123;
 GlyGlyGlySerGlyGlyGlySer SEQ ID NO:124;
 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySer SEQ ID
 NO:125;
 SerGlyGlySerGlyGlySer SEQ ID NO:126;
 15 GluPheGlyAsnMet SEQ ID NO:127;
 GluPheGlyGlyAsnMet SEQ ID NO:128;
 GluPheGlyGlyAsnGlyGlyAsnMet SEQ ID NO:129; and
 GlyGlySerAspMetAlaGly SEQ ID NO:130.

- 20 3. The EPO receptor agonist polypeptide of claim 1 selected from the group consisting of;
 SEQ ID NO:1; SEQ ID NO:2; SEQ ID NO:3; SEQ ID
 NO:4; SEQ ID NO:5; SEQ ID NO:6; SEQ ID NO:7;
 SEQ ID NO:8; SEQ ID NO:9; SEQ ID NO:10; SEQ ID
 25 NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID
 NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID
 NO:17; SEQ ID NO:18; SEQ ID NO:19; SEQ ID
 NO:20; SEQ ID NO:21; SEQ ID NO:22; SEQ ID
 NO:23; SEQ ID NO:24; SEQ ID NO:25; SEQ ID

95

NO:26; SEQ ID NO:27; SEQ ID NO:28; SEQ ID
 NO:29; SEQ ID NO:30; SEQ ID NO:31; SEQ ID
 NO:32; SEQ ID NO:33; SEQ ID NO:34; SEQ ID
 NO:35; SEQ ID NO:36; SEQ ID NO:37; SEQ ID
 5 NO:38; SEQ ID NO:39; SEQ ID NO:40; SEQ ID
 NO:41; SEQ ID NO:42; SEQ ID NO:43; SEQ ID
 NO:44; SEQ ID NO:45; SEQ ID NO:46; SEQ ID
 NO:47; SEQ ID NO:48; SEQ ID NO:49; SEQ ID
 10 NO:50; SEQ ID NO:51; SEQ ID NO:52; SEQ ID
 NO:53; SEQ ID NO:54; SEQ ID NO:55; SEQ ID
 NO:56; SEQ ID NO:57; SEQ ID NO:58; SEQ ID
 NO:59 and SEQ ID NO:122.

4. The EPO receptor agonist polypeptide of
 15 claim 3 wherein the linker sequence (GlyGlyGlyGlySer
 SEQ ID NO:123) is replaced by a linker sequence
 selected from the group consisting of;

GlyGlyGlySerGlyGlyGlySer SEQ ID NO:124;
 20 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySer SEQ ID
 NO:125;
 SerGlyGlySerGlyGlySer SEQ ID NO:126;
 GluPheGlyAsnMet SEQ ID NO:127;
 GluPheGlyGlyAsnMet SEQ ID NO:128;
 25 GluPheGlyGlyAsnGlyGlyAsnMet SEQ ID NO:129; and
 GlyGlySerAspMetAlaGly SEQ ID NO:130.

5. A nucleic acid molecule comprising a DNA
 sequence encoding the EPO receptor agonist
 30 polypeptide of claim 1.

6. A nucleic acid molecule comprising a DNA
 sequence encoding the EPO receptor agonist
 polypeptide of claim 2.

35

96

7. A nucleic acid molecule comprising a DNA sequence encoding the EPO receptor agonist polypeptide of claim 3.

5 8. A nucleic acid molecule comprising a DNA sequence encoding the EPO receptor agonist polypeptide of claim 3 selected from the group consisting of;

10 SEQ ID NO:60; SEQ ID NO:61; SEQ ID NO:62; SEQ ID NO:63; SEQ ID NO:64; SEQ ID NO:65; SEQ ID NO:66; SEQ ID NO:67; SEQ ID NO:68; SEQ ID NO:69; SEQ ID NO:70; SEQ ID NO:71; SEQ ID NO:72; SEQ ID NO:73; SEQ ID NO:74; SEQ ID NO:75; SEQ ID NO:76; SEQ ID NO:77; SEQ ID NO:78; SEQ ID NO:79; SEQ ID NO:80; SEQ ID NO:81; SEQ ID NO:82; SEQ ID NO:83; SEQ ID NO:84; SEQ ID NO:85; SEQ ID NO:86; SEQ ID NO:87; SEQ ID NO:88; SEQ ID NO:89; SEQ ID NO:90; SEQ ID NO:91; SEQ ID NO:92; SEQ ID NO:93; SEQ ID NO:94; SEQ ID NO:95; SEQ ID NO:96; SEQ ID NO:97; SEQ ID NO:98; SEQ ID NO:99; SEQ ID NO:100; SEQ ID NO:101; SEQ ID NO:102; SEQ ID NO:103; SEQ ID NO:104; SEQ ID NO:105; SEQ ID NO:106; SEQ ID NO:107; SEQ ID NO:108; SEQ ID NO:109; SEQ ID NO:110; SEQ ID NO:111; SEQ ID NO:112; SEQ ID NO:113; SEQ ID NO:114; SEQ ID NO:115; SEQ ID NO:116; SEQ ID NO:117; SEQ ID NO:118 and SEQ ID NO:119.

30 9. A nucleic acid molecule comprising a DNA sequence encoding the EPO receptor agonist polypeptide of claim 4.

10. A method of producing a EPO receptor agonist polypeptide comprising: growing under
35 suitable nutrient conditions, a host cell transformed or transfected with a replicable vector comprising said nucleic acid molecule of claim 5, 6, 7, 8 or 9

97

in a manner allowing expression of said EPO receptor agonist polypeptide and recovering said EPO receptor agonist polypeptide.

5 11. A composition comprising; a EPO receptor agonist polypeptide according to claim 1, 2, 3 or 4; and a pharmaceutically acceptable carrier.

12. A composition comprising; a EPO receptor
10 agonist polypeptide according to claim 1, 2, 3 or 4; a factor; and a pharmaceutically acceptable carrier.

13. The composition of claim 12 wherein said
factor is selected from the group consisting of: GM-
15 CSF, G-CSF, c-mpl ligand, M-CSF, IL-1, IL-4, IL-2,
IL-3, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-
12, IL-13, IL-15, LIF, flt3/flk2 ligand, human growth
hormone, B-cell growth factor, B-cell differentiation
factor, eosinophil differentiation factor and stem
20 cell factor, IL-3 variants, fusion proteins, G-CSF
receptor agonists, c-mpl receptor agonists, IL-3
receptor agonists, multi-functional receptor
agonists.

25 14. A method of stimulating the production of hematopoietic cells in a patient comprising the step of; administering a EPO receptor agonist polypeptide of claim 1, 2, 3 or 4, to said patient.

30 15. A method for selective ex vivo expansion of erythroid progenitors, comprising the steps of;
 (a) culturing erythroid progenitor cells in a culture medium, comprising; a polypeptide of claim 1, 2, 3 or 4; and
35 (b) harvesting said cultured cells.

98

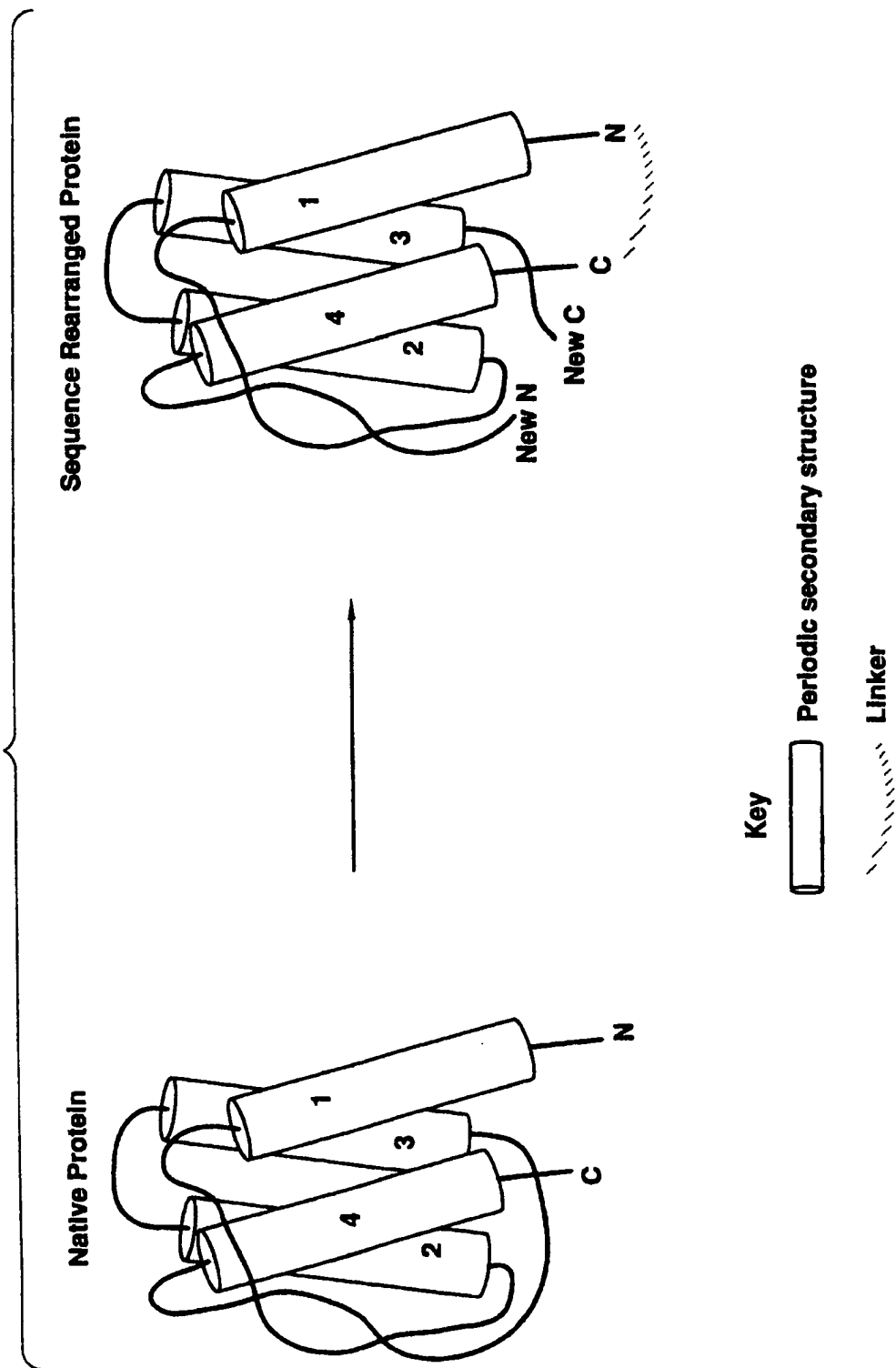
16. A method for selective ex vivo expansion of erythroid progenitors, comprising the steps of;
- (a) separating erythroid progenitor cells from other cells;
 - 5 (b) culturing said separated erythroid progenitor cells with a selected culture medium comprising a polypeptide of claim 1, 2, 3 or 4; and
 - (c) harvesting said cultured cells.
- 10 17. A method for treatment of a patient having a hematopoietic disorder, comprising the steps of;
- (a) removing erythroid progenitor cells;
 - (b) culturing said erythroid progenitor cells in a culture medium, comprising; a polypeptide of claim
 - 15 1, 2, 3 or 4;
 - (c) harvesting said cultured cells; and
 - (d) transplanting said cultured cells into said patient.
- 20 18. A method for treatment of a patient having a hematopoietic disorder, comprising the steps of;
- (a) removing erythroid progenitor cells;
 - (b) separating erythroid progenitor cells from other cells;
 - 25 (c) culturing said separated erythroid progenitor cells with a selected culture medium comprising a polypeptide of claim 1, 2, 3 or 4;
 - (d) harvesting said cultured cells; and
 - (e) transplanting said cultured cells into said
 - 30 patient.
19. A method of claim 15 wherein said erythroid progenitor cells are isolated from peripheral blood.
- 35 20. A method of claim 16 wherein said erythroid progenitor cells are isolated from peripheral blood.

21. A method of claim ⁹⁴17 wherein said erythroid
progenitor cells are isolated from peripheral blood.

22. A method of claim 18 wherein said erythroid
5 progenitor cells are isolated from peripheral blood.

1 / 6

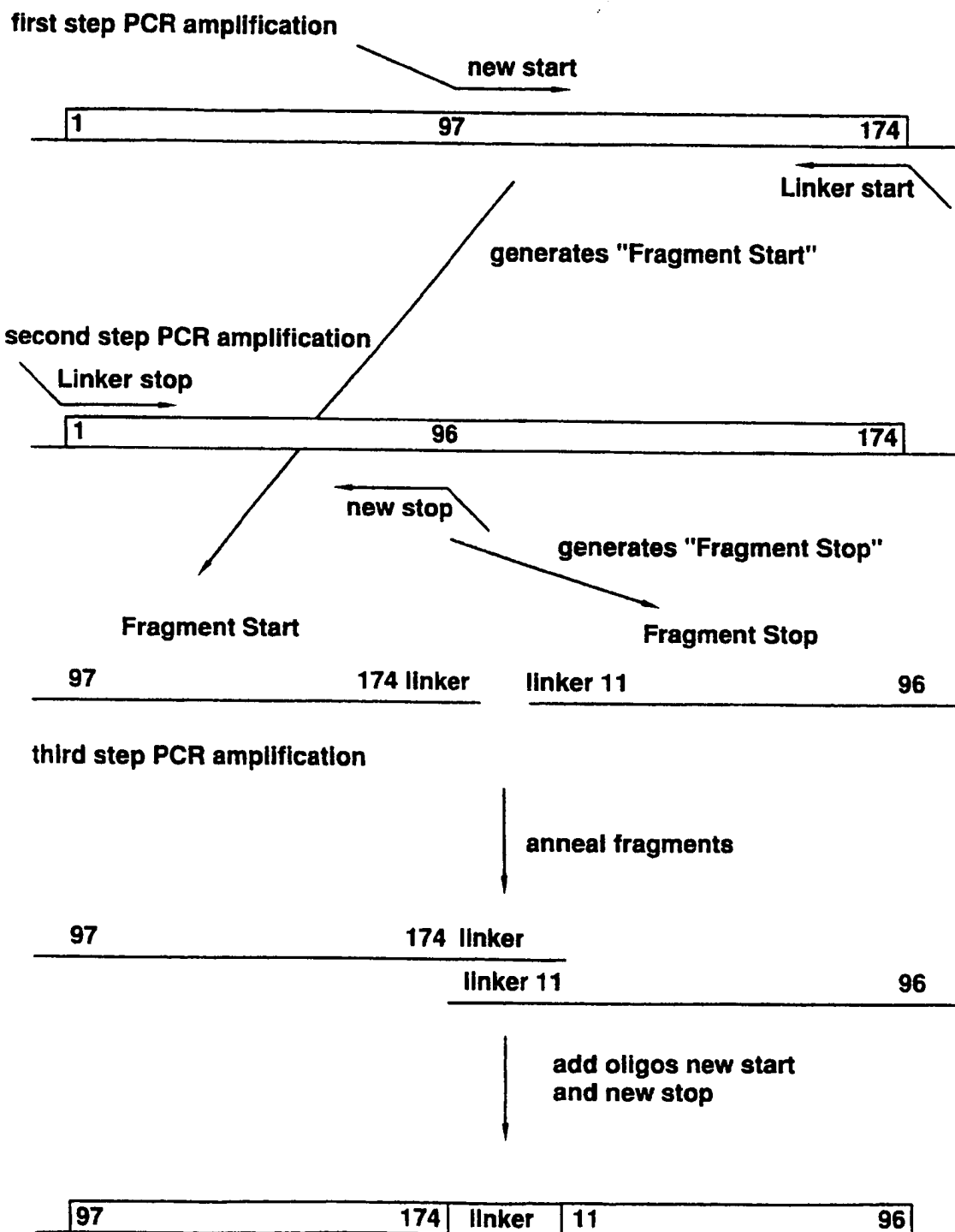
FIG.1



SUBSTITUTE SHEET (RULE 26)

2 / 6

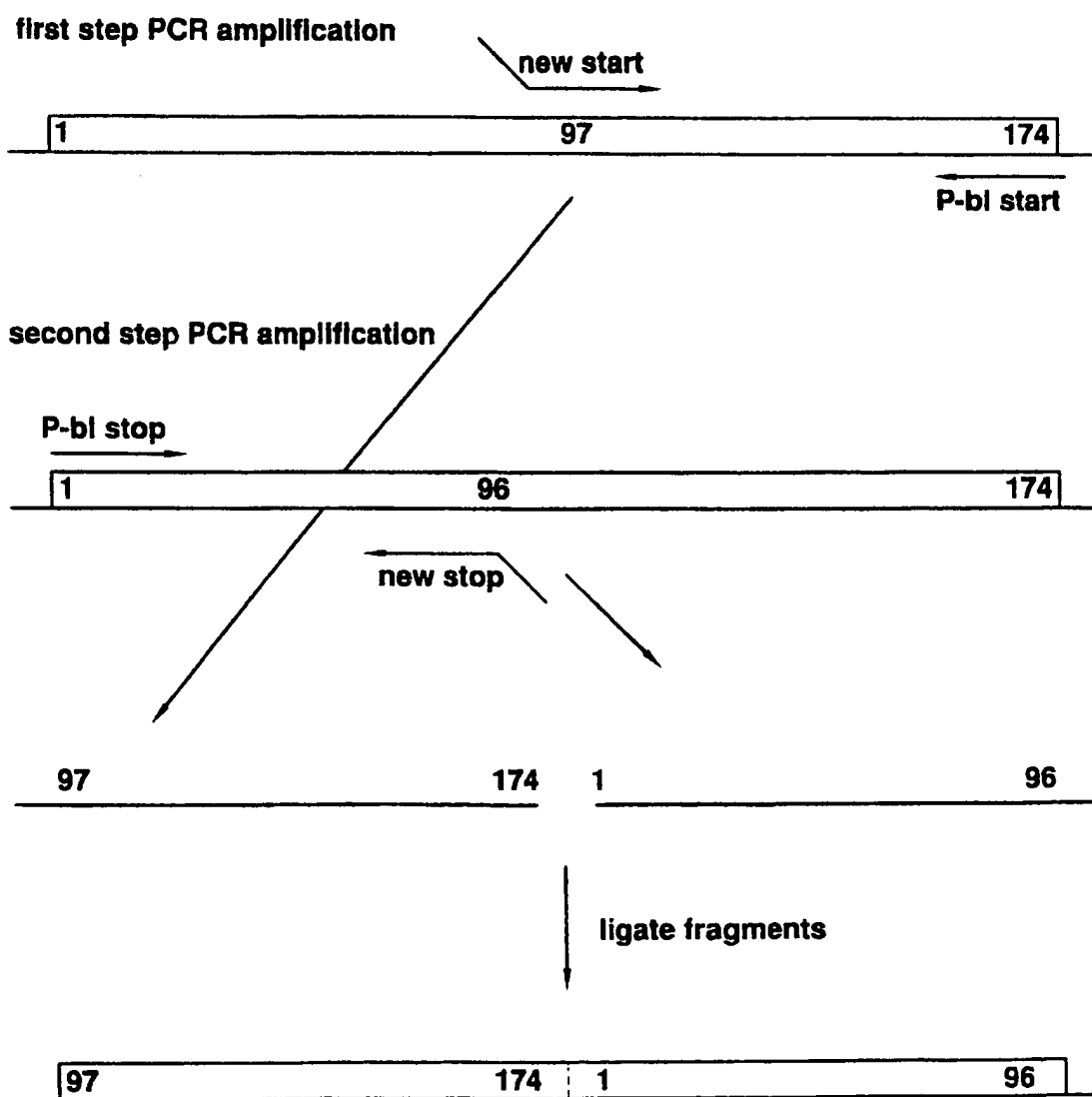
FIG.2



SUBSTITUTE SHEET (RULE 26)

3 / 6

FIG.3



4 / 6

FIG.4

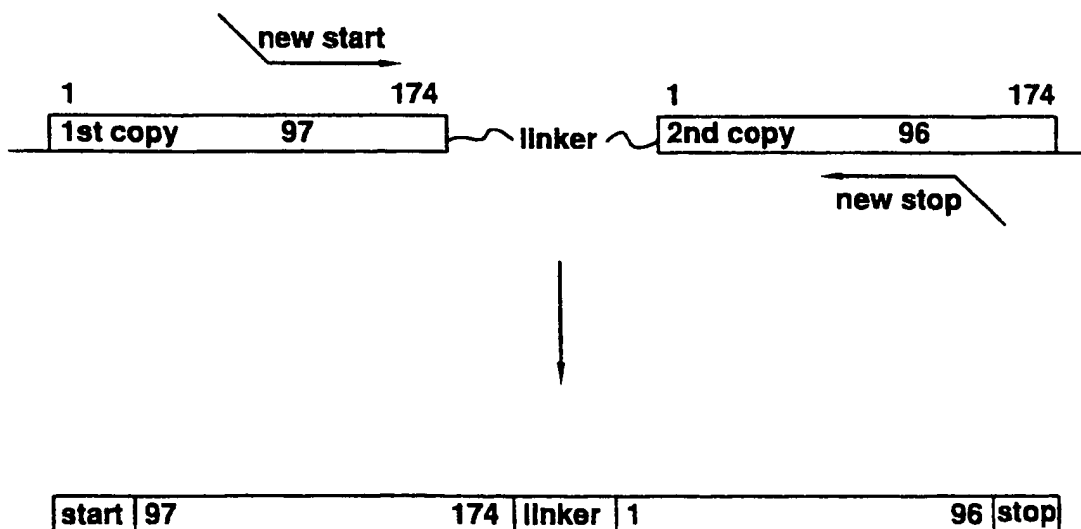
I. Construct tandemly-duplicated template**II. PCR-amplify tandemly-duplicated template**

FIG. 5A

1 G C C C C A C C A G C C T C A T C T G T G A C A G C C G A G T C C T G G A G A G G T A C C T C T T G G A G G C C C A A G
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 60 C G G G T G G T G C G G A G T A G A C A C T G T C G G C T C A G G A C C T C T C C A T G G A G A A C C T C C G G T T C
 A l a P r o A r g L e u I l e C y s A s p S e r A r g V a l L e u G l u A r g T y r L e u L e u G l u A l a L y s

 61 G A G C C G A G A A T A T C A C G A C G G G C T G T G C T G A C A C T G C A G C T T G A A T G A G A A T A T C A C T
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 120 C T C C G G C T C T T A T A G T G C T G C C C G A C A C G A C T T G T G A C G T C G A A C T T A C T C T T A T A G T G A
 G l u A l a G l u A s n I l e T h r T h r G l y C y s A l a G l u H i s C y s S e r L e u A s n G l u A s n I l e T h r

 121 G T C C C A G A C A C C A A A G T T A T T C T A T G C C T G G A A G A G A T G G A G G T C G G G C A G C A G G C C
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 180 C A G G G T C T G T G G T T C A A T T A A G A T A C G G A C C T T C T C C T A C C T C C A G C C C G T C G T C C G G
 V a l P r o A s p T h r L y s V a l A s n P h e T y r A l a T r p L y s A r g M e t G l u V a l G l y G l n G l n A l a

 181 G T A G A A G T C T G G C A G G C C T G G C C C T G C T G T C G G A A G C T G T C C T G C G G G C C A G G C C C T G
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 240 C A T C T T C A G A C C G T C C C G A C C G G A C G A C A G C C T T C G A C A G G A C G C C C C G T C C G G G A C
 V a l G l u V a l T r p G l n G l y L e u A l a L e u L e u S e r G l u A l a V a l L e u A r g G l y G l n A l a L e u

 241 T T G G T C A A C T C T T C C C A G C C G T G G G A G C C C C T G C A G C T G C A T G T G G A T A A A G C C G T C A G T
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 300 A A C C A G T T G A G A A G G T C G G C A C C C T C G G G A C G T C G A C G T A C A C C T A T T T C G G C A G T C A
 L e u V a l A s n S e r S e r G l n P r o T r p G l u P r o L e u G l n L e u H i s V a l A s p L y s A l a V a l S e r

5 / 6

6 / 6

FIG. 5B

```

301  GGCCTTCGCAGCCCTCACCACCTCTGCTTCGGGCTCTGGGAGCCCAGAAAGCCATCTCC
      +-----+-----+-----+-----+-----+-----+
      +-----+-----+-----+-----+-----+-----+
360  CCGGAAGCGTCGGAGTGGTAGACGAAGCCCGAGACCCCTCGGGTCTTCCTTCGGTAGAGG
      +-----+-----+-----+-----+-----+-----+
      GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer

361  CCTCCAGATGCGGCCCTCAGCTGCTCCACTCCGAACAATCACTGCTGACACTTTCGCAAA
      +-----+-----+-----+-----+-----+-----+
      +-----+-----+-----+-----+-----+-----+
420  GGAGGTCTACGCCGGAGTCGACGAGGTGAGGCTTGTTAGTGACGACTGTGAAAGGC GTT
      +-----+-----+-----+-----+-----+-----+
      ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys

421  CTCTTCCGAGTCTACTCCAATTTCTCTCCGGGAAAGCTGAAGCTGTACACAGGGAGGCC
      +-----+-----+-----+-----+-----+-----+
      +-----+-----+-----+-----+-----+-----+
480  GAGAAGGCTCAGATGAGGTTAAAGGAGGCCCTTTTCGACTTCGACATGTGTCCCTCCGG
      +-----+-----+-----+-----+-----+-----+
      LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla

      TGCAGGACAGGGGACAGATGA
      +-----+-----+-----+-----+-----+-----+
481  ACGTCTGTCCCTGTCTACT
      +-----+-----+-----+-----+-----+-----+
      CysArgThrGlyAspArg

```

INTERNATIONAL SEARCH REPORT

Internat	Application No
PCT/US 97/18703	

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 C12N15/18 C07K14/505 C07K14/52 A61K38/18 C12N5/10
 C12N5/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 95 27732 A (US HEALTH ; PASTAN IRA (US); KREITMAN ROBERT J (US)) 19 October 1995 see abstract; claims 1-51; figures SEQ.54-57 ---	1-13, 15, 16, 19-22
Y	WO 92 06116 A (ORTHO PHARMA CORP) 16 April 1992 see page 2, paragraph 3; claims 1-26; figure SEQ.3 ---	1-13, 15, 16, 19-22
A	VIGUERA AR ET AL: "The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics." J MOL BIOL, APR 7 1995, 247 (4) P670-81, ENGLAND, XP002056595 cited in the application see the whole document ---	1-11
-/-		

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

23 February 1998

Date of mailing of the international search report

1 1. 03 98

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Gurdjian, D

INTERNATIONAL SEARCH REPORT

Internat	Application No
PCT/US 97/18703	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>HORLICK R A ET AL: "PERMUTEINS OF INTERLEUKIN 1 BETA-A SIMPLIFIED APPROACH FOR THE CONSTRUCTION OF PERMUTATED PROTEINS HAVING NEW TERMINI" PROTEIN ENGINEERING, vol. 5, no. 5, 1992, pages 427-431, XP002022097 see the whole document</p> <p style="text-align: center;">---</p>	1-13
A	<p>KREITMAN R J ET AL: "A CIRCULARLY PERMUTED RECOMBINANT INTERLEUKIN 4 TOXIN WITH INCREASED ACTIVITY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 91, no. 15, July 1994, pages 6889-6893, XP002022099 see the whole document</p> <p style="text-align: center;">---</p>	1-13
A	<p>WO 95 21197 A (SEARLE & CO ;BAUER CHRISTOPHER S (US); ABRAMS MARK ALLEN (US); BRA) 10 August 1995 see page 1 - page 33</p> <p style="text-align: center;">-----</p>	1-13, 15, 16, 19-22

INTERNATIONAL SEARCH REPORT

Int. application No.
PCT/US 97/18703

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 97/18703

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Remark : Although claims 14 17 18 are directed to a method of treatment of the human/animal body , the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 97/18703

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9527732 A	19-10-95	US 5635599 A	03-06-97
		AU 2285795 A	30-10-95
		CA 2187283 A	19-10-95
		EP 0754192 A	22-01-97

WO 9206116 A	16-04-92	AU 1157695 A	13-04-95
		AU 8735991 A	28-04-92
		CA 2069746 A	29-03-92
		EP 0503050 A	16-09-92
		JP 5502463 T	28-04-93
		ZA 9107766 A	29-03-93

WO 9521197 A	10-08-95	AU 1680595 A	21-08-95
		EP 0742796 A	20-11-96
		JP 9508524 T	02-09-97
